

# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 189439**

**TO: Andrew D Kosar**  
**Location: 3c04 / 3c18**  
**Art Unit: 1654**  
**Wednesday, May 10, 2006**

**Case Serial Number: 10/800179**

**From: Noble Jarrell**  
**Location: Biotech-Chem Library**  
**Rem 1B71**  
**Phone: 272-2556**

**Noble.jarrell@uspto.gov**

### **Search Notes**

This Page Blank (uspto)

189439

STIC-Biotech/ChemLib

From: ANDREW KOSAR [andrew.kosar@uspto.gov]  
Sent: Tuesday, May 09, 2006 5:16 PM  
To: STIC-Biotech/ChemLib  
Subject: Database Search Request, Serial Number: 10/800,179

Requester:  
ANDREW KOSAR (P/1654)  
Art Unit:  
GROUP ART UNIT 1654  
Employee Number:  
80341  
Office Location:  
REM 03C04  
Phone Number:  
(571)272-0913  
Mailbox Number:  
REM 3c18

Case serial number:  
10/800,179  
Class / Subclass(es):

Earliest Priority Filing Date:

Format preferred for results:  
Paper

Search Topic Information:  
Please search SEQ ID NO:19 in 10/800,179.


Special Instructions and Other Comments:

\*\*\*\*\*  
Searcher: Noble  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: 5/12/06  
Searcher Prep Time: 5  
Online Time: 5

\*\*\*\*\*  
Type of Search  
NA# \_\_\_\_\_ AA#: 1  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*  
Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: Compugen  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_

This Page Blank (uspto)

 **PALM INTRANET**

Day : Wednesday

Date: 5/10/2006

Time: 16:58:59

## Inventor Information for 10/800179

Inventor Name	City	State/Country
KUMAR, MANOJ	FREMONT	CALIFORNIA
CUEVAS, WILLIAM A.	SAN FRANCISCO	CALIFORNIA

[Appln Info](#)[Contents](#)[Petition Info](#)[Atty/Agent Info](#)[Continuity Data](#)[Foreign Data](#)[Inventor](#)Search Another: Application#   or Patent#  PCT /  /   or PG PUBS #  Attorney Docket #  Bar Code #  

To go back use Back button on your browser toolbar.

Back to [PALM](#) | [ASSIGNMENT](#) | [OASIS](#) | [Home page](#)

**This Page Blank (uspto)**

## Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

**This Page Blank (uspto)**



November 2005

Published\_Applications Nucleic Acid and Published\_Applications Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases; older published applications make up the Published\_Applications\_Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New).

Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).

This Page Blank (uspto)

This Page Blank (uspto)







A:	Accession:	B26728
A:	Molecule type:	mRNA
A:	Residues:	1,'RS','4-11','E','13-225,240-636','V','638-747'<RA>
A:	Cross-references:	UNIPARC:UPI000002A8D3; GB:K03505; NID:9163025; PIDN:AAA30505.1; PID:
A:	Accession:	C26728
A:	Molecule type:	mRNA
A:	Residues:	1,'RS','4-11','E','13-225,260-636','V','638-747'<RA>
A:	Cross-references:	UNIPARC:UPI0000002A8D; GB:K03506; NID:9163027; PIDN:AAA30506.1; PID:
R:	Citella, G.; May, M.; Ornstein-Goldstein, N.; Indik, Z.; Morrow, S.; Yen, H.S.; Rosenberg,	
A:	Title:	Structure of the 3' portion of the bovine elastin gene.
A:	Reference number:	A22343; PMID:85280426; PMID:2992576
A:	Accession:	A22343
A:	Molecule type:	DNA
A:	Residues:	613-747<CIC>
A:	Cross-references:	UNIPARC:UPI0000173CAF; GB:M20415
R:	Rosenblum, J.	
Lab:	Invest.	51, 605-623, 1984
A:	Title:	Biology of disease: Elastin: Relation of protein and gene structure to disease.
A:	Reference number:	I45885; PMID:8505254; PMID:6150137
A:	Accession:	I45886
A:	Status:	preliminary; translated from GB/EMBL/DDBJ
A:	Molecule type:	DNA
A:	Residues:	678-683, 685-747<ROS>
A:	Cross-references:	UNIPARC:UPI000008A56A; GB:M31898; NID:9163015; PIDN:AAA96417.1; PID:
R:	Brown, P.L.; Mecham, L.; Tisdale, C.; Mecham, R.P.	
Biochem.	Biophys. Res. Commun.	186, 549-555, 1992
A:	Title:	The cysteine residues in the carboxy terminal domain of tropoelastin form an intramolecular disulfide bond
A:	Reference number:	A58621; PMID:92337651; PMID:1633791
A:	Contents:	annotation, disulfide bonds
A:	Comment:	The term tropoelastin refers to a soluble precursor form of the extracellular matrix protein elastin.
C:	Genetics:	
A:	Introns:	634/3; 653/3; 676/3; 689/3; 707/3; 716/3; 733/3
A:	Note:	the list of introns is incomplete
C:	Superfamily:	elastin
C:	Keywords:	alternative splicing; extracellular matrix; glycoprotein; hydroxyllysine
F:	1-747/Product:	elastin precursor, splice form a #status predicted <EPA>
F:	1-225,260-747/Product:	elastin precursor, splice form c #status predicted <EPC>
F:	1-225,240-747/Product:	elastin precursor, splice form b #status predicted <EPB>
F:	1-26/Domains:	signal sequence #status predicted <SIG>
F:	1-27-747/Product:	elastin #status predicted <MAT>
F:	105,109,252,271,275,324,327,400,404,407,445,448,489,493,544,548,552,606,609,645,649,688,737-742/Disulfide bonds:	#status experimental

Query Match	27.6%	Score 1151;	DB 1;	Length 747;
Best Local Similarity	41.4%;	Pred. No. 1.1e-62;		
Matches	350;	Conservative 46;	Mismatches 249;	Indels 200; Gaps 45;
OY	5	GSGAGGSGGV- -GVGPVGVPKGVPVGPGVPG- - - - -	VGPVG- - - - -	PGA 50
Db	27	GGVPGAVP- - -GVPGVFPFGAGLGLGVGLGFPGVKPKAPGVGILVGPLGABGSLPPGA 85		
OY	51	GAGSGAGGSG- -AGAGSGAGSGGVPGVPGVPGVPGKGVPGVPGVPGVPGVPGVPG 109		
Db	86	PFQGFAGGGAAGAAYAAAKAAGAAAGLVGGIG- - - - -GVGGLGVSTG- - - - -AVPQG 136		
OY	110	AGAGSGAGSGAGSGAGSGAGSGGVPGVPGVPGVPGKGVPGV- - -PGVPGVPGV 166		
Db	137	LQAGVAGVKK- - - - -VPGVLPGV- - -YRGGVLPFGAARPGI- - -GVLPGV 180		
OY	167	PGAGAGSGAGSGAGSGAGSGAGSGGVPGVPGVPGVPGV- - -KGVPGVPGVPGV- - - 223		
Db	181	PLGAGVKKPRQVAGARA- - - - -GRFVGFPGQQPLPLGYPIKAPRLPYGTGL 230		
OY	224	- - - - -PGVPGAGSGAGSGAGSG- - - - -AGAGSGVPGVPGV- - -PGVPGV 267		
Db	231	PYKTGLPVFGPGGVASGAKAGYPTGYGVPGQAAAAAKAAKLGAAGAGVLPVGVG 290		
OY	268	GCGVPGVPGVPGVPGVPGVPGVPGAGAGSGAGSGAGSGAGSGAGSGAGSGVPGVPGVPGV 327		
Db	291	GRPIPG- - - - -ACGALPGIGLAGVGAADAAAAAATAKAATFGAAGGILPGVPGVPGV 345		

[illegible]



[illegible]

Db 1274 GADGNAAGRGACVPDPDGDGCNAAFGAATPGVGGCIGTGTGAGAAGACA 1322

RESULT 8  
B70807  
hypothetical glycine-rich protein RV3512 - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: B70807  
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
J.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 383, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; PMID:98295987; PMID:9634230  
A:Accession: B70807  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1079 <COL>  
A:Cross-references: UNIPARC:UPI00000D3AEA; GB:A.L022022; GB:AL123456; NID:gj261554; PIDN:(C:  
C:Genetics:  
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

Query Match 26.1%; Score 1089.5; DB 2; Length 1079;  
Best Local Similarity 36.6%; Pred. No. 6,9e-59;  
Matches 317; Conservative 32; Mismatches 394; Indels 123; Gaps 31,

QY 1 GAGAGSAGAGSGGVGPVGVPFVGPKGVPEVGGVPPGVPPGVPGA--GAGSAG- 57  
DB 242 GAGDGGHGGTGTAAGGNGCTG---GAGCGSIDGVGGGTG-GTGNNGNGAIIGAAGDAAG 296  
QY 58 ----AGSGAASGAGSGAGSGGVGPVGVPFVGPKGVPEVGGVPPGVPPGAAG 113  
DB 297 SGNSGNGITGKKGNNAGAG-GAAGSNGTGANNTGDDGNCAGAAATRASNGAGGTG 355  
QY 114 S-GAGAGSAGAGSGAGSGAGSGGVPPGVPPGVPPGK-----GVPPGVPPG----V 162  
DB 356 SAAGNGCTGRRRGSSGAGGDTIGTVGCKGKNCADEVGAGAGGS GFPTSPGNGGG 415  
QY 163 GPVGPPAGAGSAGAGSAGAGSGAGSGAGSGGVPPGVPPGV-----GVPPGVP 212  
DB 416 CGGSGGAGAGAGAG-GAGGAGANTGANGGQGAGGTGGAASASANNGSCGAGTGD 474  
QY 213 GVGPVGPVGPVGVPVGAAGAGSG---AGAGSAGAGSAGAGSGGVPPGVPPGVVP 267  
DB 475 GSGGAGH-GTGAAGTTGAAGDGGCGGCGGAGGAGGCGGAGAGGTGNGCNTTGTATG 533  
QY 268 GKCPVGPVGPVGPVGPVPGAAGSAGAGSAGAGSAGAGSGGVPPGVPPGVVP 327  
DB 534 TAGAAGNGAANKG---GAGGCGCTGGTGGCGGAGGADGAGGTGGDRVTYGAGTVPAGSG 590  
QY 328 GKCPVGPVGPVGPVGPVPGAAGSAGAGSAGAGSAGAGSGGVPPGVPPGVVP 387  
DB 591 CGGAGNAGGAGAG---GCGGADGGSGDGDPACTGCGNAGCNRNNSGVTGAGAGNGGGA 647  
QY 388 GKCPVGPVGPVGPVGPVPGAAGSAGAGSAGAGSAGAGSGGVPPGVPPGVVP 447  
DB 648 NGAGAGAG-GSGGGTCGNGAGGADADAGNGAGNGTGNAGNGNG-----GIAGNG-- 698  
QY 448 GKCPVGPVGPVGPVGPVPGAAGSAGAGSAGAGSAGAGSGGVPPGVPPGVVP 507  
DB 699 GNGAGTAGTSAGNAGNGSG--GNGNAGMGINSIGSGEDGAGAGNGGAAGTGTGTDG--GLT 756  
QY 508 GKCPVGPVGPV-----GPVGPVGPVPGAAGSG---AGA 538  
DB 757 GTGTGTGSGGTGGDGGNGAGDNMTANMATAAGAGDGGNGGDDGFGGAGAGGGGLTAGAN 816  
QY 539 -----GSAAGAGSGA-----GAGSGVGVPPGVPPGVPPGVPPG 577  
DB 817 GTGGGGAAGDGGNAGIIGHGRLTDPGANGCTGGAGNGCTGTGAGAGTGSLSGGGTGGGAGN 876









Db 410 GKGAGNGAGAGAGLVTNGAG -GAGNGAPGAPPSGDPNCGGGAGAGGKGDGA 468  
Qy 241 GAGAGGAGAGSGVPGVPGVPGKGVPGVPGV -GPGVPGVPGAGAGSGAG 239  
Db 469 QAGDGGAGAGKGGKNGATG---ATGTLNGLAGADDTGGKKGNGAGGGAGGQ 524  
Qy 300 SG-----AGAGSGAGSGVPGVPGVPGV-----GVPGKGVPGVPGV 339  
Db 525 GGTALAAATTODSGMAGAGAGAGMGDGGNGAKTPTDNGDDVCGAGNGAGSRGIG 584  
Qy 340 --PGVPGVPGAGAG-----GSGAGAGSGA-----GAGSGAGAGSGVPGV 380  
Db 585 GAGGIG-GAGSTAGAGARCATPTSGNGGTGAGANAATVACGAGAGKGGKNGGLVNG 643  
Qy 381 VPGVPGVPGKGVPGVPGV-----GVPGVPGAGAGAGAGAGAGAGAG 431  
Db 644 GAG-GKGGGMAVA--GSSPTTAGESGTSQNGAG-GAGAGAGRGDFGGDGTGAGG 700  
Qy 432 SGVPGVPGVPGV--GVPGKGVPGVPGVPGV-----GVPGAGAGAG 475  
Db 701 NGANGANAATTPGAKGDDGGHGPAGAGNGGGGPGGLAGNLFGQNGIGVG--GSGAG 759  
Qy 476 AG-----AGSGA-----GAGSGAGAGSGVPGVPGV-----VGPG-KGVPGVPGV 518  
Db 760 AGGLAGDGGNGAGNFAFDPDNGAGHGGNG-GNPGAGGGGSGAGSTPGAKAHFTPTS 818  
Qy 519 GPGVPGV-----VPGAGAGSGAGAGAGAGAGAGSGVPGVPGV---PGV 566  
Db 819 GGGGGGAGNGNSQVVGAGNGDGGNGNGAGTGTGNGRGGDARFGAGSANAATNPENG 878  
Qy 567 PGKGVPGVPGVPGVPGVPGAGAGSGAGAGAGAGAGAGAGSGVPGVPGV 626  
Db 879 P-NGNPGAG 935  
Qy 627 PGKGVPGVPGVPGVPGV-----GPGAGAGSGAGAGAGAGAGAGAGAG 674  
Db 936 GGGAG 995  
Qy 675 GVPVPGVPGVPGVPGVPGV-----VGP-----GVPGVPGVPGAGAGAG 715  
Db 996 GGGGAG 1055  
Qy 716 AGAGSGAG 755  
Db 1056 GNGGTGAG 1114  
Qy 756 -----PGV-----PGVPGVPGAGAGAGAGAG 779  
Db 1115 GNGGTGTGAG 1160

## RESULT 14

E70806  
hypothetical glycine-rich protein Rv3507 - Mycobacterium tuberculosis (strain H37RV)  
C/Species: Mycobacterium tuberculosis  
C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C/Accession: E70806  
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Comor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajadream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
A/Authors: Squares, R.; Sultson, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A/Reference number: A70500; MIMD:9825987; PMID:9654230  
A/Accession: E70806  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-1181 <COU>  
A/Cross-references: UNIPARC:UPI000003JA67; GB:AL022022; GB:AL123456; MID:93261554; PIDN:  
A/Experimental source: strain H37RV  
C/Genetics:  
A/Gene: Rv3507  
C/Superfamily: collagen alpha 1(IV) chain

Query Match 24.1%; Score 1006; DB 2; Length 1181;  
Best Local Similarity 35.1%; Pred. No. 8 4e-54;  
Matches 307; Conservative 31; Mismatches 384; Indels 152; Gaps 36;

Qy 1 GAGAGAGAGAGAG-----VGPGVPGVPGKGVPGVPGVPGVPGVPGV---VGPAG 51  
Db 169 GAGGSGAAG 224  
Qy 52 AGSGAG 110  
Db 225 GAG 282  
Qy 111 GAG-----SGAG 160  
Db 283 GAGGLPMNGDGGAG 341  
Qy 161 -----GVPGVPG-----AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 194  
Db 342 GAG 401  
Qy 195 GVPVPGVPGVPGVPGK-----VPGVPGVPGVPGVPGVPGVPGVPGVPGV 249  
Db 402 GAGGTGTG-GTGGNADAAAVVFGANDPFAGAGAGAGAGAGAGAGAGAGAGAG 460  
Qy 250 AGSGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 306  
Db 461 KG---GTGAG 512  
Qy 307 -----GAG 353  
Db 513 GGGGTGTGAG 568  
Qy 354 SGAG 413  
Db 569 TGGNAG 622  
Qy 414 SGAGAG-----GAG 468  
Db 623 NGGTAGSGGTGAG 678  
Qy 469 GAG 528  
Db 679 AAGAGGTGAG 735  
Qy 529 GAG 588  
Db 736 GGNAG 789  
Qy 589 GAG 646  
Db 790 GSGGTGAG 842  
Qy 647 GPGAGAGAGAGAGAG-----AGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 699  
Db 843 GNGAG 898  
Qy 700 PGVPGVPGAG 745  
Db 899 GNGAG 957  
Qy 746 VPGK-GVPVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 778  
Db 958 TGGNAG 986

## RESULT 15

A70934  
hypothetical glycine-rich protein Rv0578c - Mycobacterium tuberculosis (strain H37RV)  
C/Species: Mycobacterium tuberculosis  
C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C/Accession: A70934  
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrel, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: A70934  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Rebids: 1-1306 <COL>  
 A:Cross-references: UNIPROT:O53775; UNIPARC:UPI00000D4FC8; GB:AL021942; GB:AL123456; NID  
 A:Experimental source: strain H37Rv  
 C:Genetics:  
 A:Gene: Rv0578c  
 C:Superfamily: collagen alpha 1(IV) chain

Query Match 23.6%; Score 984; DB 2; Length 1306;  
 Best Local Similarity 34.0%; Pred. No. 1.7e-52;  
 Matches 332; Conservative 34; Mismatches 386; Indels 224; Gaps 44;

QY 1 GAGA-GSGAGAGSGVPGVPGVPGKGVPGV-----GPGVPGVPGVPGVPGV 50  
 DB 198 GAGAVGNGGAGGTAGLFGVGGAG-GAGGNGIAGTGTASTPGSGGTAGAG-GIGNG 255  
 QY 51 GAGS-----GAGAGSGAGSGAGAGSGVPGVPGVPGK-----GVPVGP 96  
 DB 256 GAGAGGVLMNGNGGAGGREGPGAGAGAGSAGHATNLGADGQAGNGNGAGAGGTGG 315  
 QY 97 GVGPGVPGVPGAGAGAGAGAGAGAGAGSGAGSGVPGVPGVPGVPGKGVPGV 156  
 DB 316 VGGFGGGHGL-LGLGSHGAGAGAGSGGDDGAPGDDGNGATGTWGHNLGAGGTGNGGNGP 374  
 QY 157 GVGPGVPG--VPGAGAGSGA-----GAGSGAGAGSGA-----GAGSGVPGV 199  
 DB 375 GAGAGAGAGAGSAGSAGSAGHANGAPGTSTSGNGGDDGKADALISSGQTGANAGRGDDG 434  
 QY 200 GVPVPGVPGKVPVPGVPG-----PGVPGVPGAGAGAGSGAGAGSGAGAG 251  
 DB 435 QVNGAGAGAGAGGAGGAGLFGSEAPRPGAGAGTG-GAGNGGTQAGDG-GTGAGAGAG 492  
 QY 252 SGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 307  
 DB 493 GDGSSGAGAGTGFNASSAPGAAGSPGNGNGNGPG--GAGBEGAGGLALASSGNGSG 549  
 QY 308 AGAGSGVPGVPGVPGVPG--KGVPGVPGV-----PGVPGVPGAGAGAGAG 358  
 DB 550 AGDGGAGAGNG-GTPGNGHGAALGVNGGVGAGGHGDDPGVG-GAGGQGSSTPGA 607  
 QY 359 -----GSGAGAGSGA-----GAGSGVPG--VGPVPGVPG--KVPVPG 393  
 DB 608 NGABGNTPTSGGNGNGRGADATGFGQTGASGGRGDDGLVNGAGAGAGNGSKLPG 667  
 QY 394 VGPVPGVPGVPGVPGAGAGAGAGAGAGAGAGSGV-GVPVPGVPGV----- 444  
 DB 668 LGRIGNPGILDGTGNGAGAGSGGAMAGNGTGGAGGTGGTGSSSDVNGSSAGAD 727  
 QY 445 GVPG--KGVPGVPGVPGVPGVPGAGAG--GAGAGAGAGAGAGAGSGVPGV-- 499  
 DB 728 GHPGOTGGVGTGGKGDGCGGAAPNGVAGSQGPGAGDGTGGVGGNGRGITDADG 787  
 QY 500 -----GVPVPGVPGKVPVPGVPGVPGVPGVPGVPGVPGVPGVPGV 549  
 DB 788 ATAGARQDDGAGAGAG--GKG--GRGTGGPG--GAGPAGTTGSGAGAGNGSGSGTGAD 840  
 QY 550 AGSGVPGVPGVPGVPGKVPVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 598  
 DB 841 PGDG-----GNGANGSVFTNNNGIANG--GNGNAPSGAGSGAGSTFGATGSSSIHYN 895  
 QY 599 -----GSGAGAGSGAGAGSGV-----GVPVPGVPGVPGKGVPGVPGV 638  
 DB 896 GGNGGNGGNDHALSNGAGAGNGGNGGSLRSGAGAGGHGNG--GNAASRGMGDDGCTG 954  
 QY 639 GPGVPG--VPGAGAGAGAGAGAGAGAGSG--VGPVPGVPGVPGKGVPGV 695

DB 955 GAGNAGQITGNGAGAGNGDGTGSDGNPGLTTSGGKGGDGGVGGGSSVAGDDAGGR 1014  
 QY 696 PGGVPGVPGV-----GPGAGAGSGAGAGSGA----- 722  
 DB 1015 CGAGGTGTLRGTGTGATGATGTFDAGADGHGNGGTGGVGGTGGAGGGGNGAGKAL 1074  
 QY 723 -----GAGSGAGAGSGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 764  
 DB 1075 SPTGNGSQAGAGDGAAGAGGTGGTGDG---GRGAHGTLLSSLAGTGTGNGGTG- 1129  
 QY 765 GVPAGAGAGAGAGS 780  
 DB 1130 GTGCTGAGAGAGGTGS 1145

Search completed: May 10, 2006, 10:44:30  
 Job time : 51 secs

This Page Blank (uspt)

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2006, 10:36:23 ; Search time 232 Seconds  
(without alignments)  
2372.035 Million cell updates/sec

Title: US-10-800-179-19

Perfect score: 4173

Sequence: 1 GAGAGSGAGAGSGVGPVG.....GVGPVGPGAGAGSGAGAGS 780

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database :

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2107.5	50.5	5263	1	FIBH_BOMMO
2	1916.5	45.9	897	2	O6Q254_AGEAP
3	1531	36.7	760	2	O6PY84_KUKAI
4	1459.5	35.0	1002	2	O9BI08_ARGTR
5	1451.5	34.8	988	2	O17434_NBPCL
6	1397.5	33.5	1553	2	O6PFR0_CANGA
7	1365	32.7	1884	2	O9NHW2_9ARAC
8	1314	31.5	750	1	ELN_CHICK
9	1312.5	31.5	1713	2	O9VIR6_DROME
10	1302.5	31.2	912	2	O9BIR2_PLETR
11	1302.5	31.2	2249	2	O9NHW4_NBPCL
12	1297.5	31.1	1953	2	O9BIR7_9ARAC
13	1274	30.5	1729	2	O9U617_DROME
14	1230.5	29.5	907	2	O44359_NBPCL
15	1219.5	29.2	860	2	O8C918_MUSCU
16	1218	29.2	860	1	ELN_MOUSE
17	1215	29.1	810	2	O9ES29_MOUSE
18	1195.5	28.6	875	2	OSRKH4_RAT
19	1185	28.4	864	2	ELN_RAT
20	1159	27.8	1071	2	O7YU48_DROME
21	1142	27.4	747	1	ELN_BOVIN
22	1137	27.2	672	1	PHXR3_MOUSE
23	1136	27.2	871	2	O44358_NBPCL
24	1100	26.4	1217	2	O8V1Y9_MYCTU
25	1096	26.3	1901	1	PG54_MYCTU
26	1094.5	26.2	992	2	O7TWB8_MYCBO
27	1089.5	26.1	1079	2	O6MMW7_MYCTU
28	1087	26.0	1938	2	O7TWC0_MYCBO
29	1086	26.0	1715	2	O8V1Z0_MYCTU
30	1084	26.0	1489	2	O6MMW6_MYCTU
31	1077	25.8	1460	2	O7TWC3_MYCBO

32	1064	25.5	1468	2	O9GUB5_GALME
33	1058.5	25.4	786	1	ELN_HUMAN
34	1041	24.9	747	1	SPDI_NBPCL
35	1034	24.8	1660	2	O79FD4_MYCTU
36	1034	24.8	1660	2	O7D721_MYCTU
37	1019	24.4	651	2	O9BI09_ARGTR
38	1017.5	24.4	854	2	O9BI04_DOLTE
39	1010.5	24.2	1329	2	O79FP2_MYCTU
40	1007.5	24.1	1408	2	O8VK17_MYCTU
41	1007	24.1	1408	2	O7U022_MYCBO
42	1006	24.1	1381	2	O6MMW9_MYCTU
43	1003	24.0	1384	2	O8V1Z1_MYCTU
44	998.5	23.9	1274	2	O6CHN8_YARLI
45	992	23.8	1360	2	O7TWC4_MYCBO

#### ALIGNMENTS

RESULT 1

ID	FIBH_BOMMO	STANDARD	PRT	5263 AA.
AC	P05790; Q17220; Q26379;			
DT	01-NOV-1988 (Rel. 09, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	13-SEP-2005 (Rel. 48, Last annotation update)			
DE	Fibroin heavy chain precursor (Fib-H) (H-fibroin).			
GN	Name=FIBH;			
OS	Bombyx mori (Silk moth).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;			
OC	Bombycidae; Bombyx.			
OX	NCBI_TaxID=7091;			
RN	[1]			
RN	NUCLEOTIDE SEQUENCE.			
RX	MEDLINE=20330362; PubMed=10871375; DOI=10.1093/nar/28.12.2413;			
RA	Zhou C.-Z., Confalonieri F., Medina N., Zivanovic Y., Esmaili C.,			
RA	Yang T., Jacquet M., Janin J., Duguet M., Perras R., Li Z.-G.;			
RT	"Fine organization of Bombyx mori fibroin heavy chain gene.";			
RL	Nucleic Acids Res. 28:2413-2419(2000).			
RN	[2]			
RN	NUCLEOTIDE SEQUENCE OF 1-168.			
RX	MEDLINE=80045039; PubMed=498286; DOI=10.1016/0092-8674(79)90075-8;			
RA	Tsuji moto Y., Suzuki Y.;			
RT	"The DNA sequence of Bombyx mori fibroin gene including the 5'			
RT	flanking, mRNA coding, entire intervening and fibroin protein coding			
RT	regions.";			
RL	Cell 18:591-600(1979).			
RN	[3]			
RN	PARTIAL NUCLEOTIDE SEQUENCE.			
RX	MEDLINE=79211211; PubMed=455439; DOI=10.1016/0092-8674(79)90018-7;			
RA	Tsuji moto Y., Suzuki Y.;			
RT	"Structural analysis of the fibroin gene at the 5' end and its			
RT	surrounding regions.";			
RL	Cell 16:425-436(1979).			
RN	[4]			
RN	PARTIAL NUCLEOTIDE SEQUENCE.			
RC	STRAIN=Kinsu X Shoua;			
RX	MEDLINE=8094868; PubMed=3210244;			
RA	Mita K., Ichimura S., Zama M., James T.C.;			
RT	"Specific codon usage pattern and its implications on the secondary			
RT	structure of silk fibroin mRNA.";			
RL	J. Mol. Biol. 203:917-925(1988).			
RN	[5]			
RN	PARTIAL NUCLEOTIDE SEQUENCE.			
RX	MEDLINE=94365842; PubMed=7916056; DOI=10.1007/BF00175878;			
RA	Mita K., Ichimura S., James T.C.;			
RT	"Highly repetitive structure and its organization of the silk fibroin			
RT	gene.";			
RL	J. Mol. Evol. 38:583-592(1994).			
RN	[6]			
RN	NUCLEOTIDE SEQUENCE OF 5179-5263, AND DISULFIDE BONDS.			
RC	STRAIN=J-139;			

RX MEDLINE=99296390; PubMed=1036722; DOI=10.1016/S0167-4838(99)00088-6;  
RA Tanaka K., Kajiyama N., Ishikura K., Waga S., Kikuchi A., Ohtomo K.,  
RA Takagi T., Mizuno S.;  
RT "Determination of the site of disulfide linkage between heavy and  
RT light chains of silk fibroin produced by Bombyx mori.";  
RN Biochim. Biophys. Acta 1432:92-103(1999).  
RM [7]  
RP PARTIAL NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=Dalco;  
RX PubMed=15951204; DOI=10.1126/science.1102210;  
RA Xia Q., Zhou Z., Lu C., Cheng D., Dai F., Li B., Zhao P., Zha X.,  
RA Cheng T., Chai C., Pan G., Xu J., Liu C., Lin Y., Qian J., Hou Y.,  
RA Wu Z., Li G., Pan M., Li C., Shen Y., Lan X., Yuan L., Li T., Xu H.,  
RA Yang G., Wan Y., Zhu Y., Yu M., Shen W., Wu D., Xiang Z., Yu J.,  
RA Wang J., Li R., Shi U., Li H., Li G., Su J., Wang X., Li G., Zhang Z.,  
RA Wu Q., Li J., Zhang Q., Wei N., Xu J., Sun H., Dong L., Liu D.,  
RA Zhao S., Zhao X., Meng Q., Lan F., Huang X., Li Y., Fang L., Li C.,  
RA Li D., Sun Y., Zhang Z., Yang Z., Huang Y., Xi Y., Qi O., He D.,  
RA Huang H., Zhang X., Wang Z., Li W., Cao Y., Yu Y., Yu H., Li J.,  
RA Ye J., Chen H., Zhou Y., Liu B., Wang J., Ye J., Ji H., Li S., Ni P.,  
RA Zhang J., Zhang Y., Zheng H., Mao B., Wang W., Ye C., Li S., Wang J.,  
RA Wong G.K.-S., Yang H.;  
RT "A draft sequence for the genome of the domesticated silkworm (*Bombyx*  
RT mori)." ;  
RL Science 306:1937-1940(2004).  
RM [8]  
RN SUBUNIT.  
RX MEDLINE=20568317; PubMed=10986287; DOI=10.1074/jbc.M006897200;  
RA Inoue S., Tanaka K., Arisaka F., Kimura S., Ohtomo K., Mizuno S.;  
RT "Silk fibroin of *Bombyx mori* is secreted, assembling a high molecular  
RT mass elementary unit consisting of H-chain, L-chain, and p25, with a  
RT 6:6:1 molar ratio." ;  
RL J. Biol. Chem. 275:40517-40528(2000).  
CC -I- FUNCTION: Core component of the silk filament; a strong, insoluble  
CC and chemically inert fiber.  
CC -I- SUBUNIT: Silk fibroin elementary unit consists in a disulfide-  
CC linked heavy and light chain and a p25 glycoprotein in molar  
CC ratios of 6:6:1. This results in a complex of approximately 2.3  
CC MDa.  
CC -I- TISSUE SPECIFICITY: Produced exclusively in the posterior (PSG)  
CC section of silk glands, which are essentially modified salivary  
CC glands.  
CC -I- DOMAIN: Composed of antiparallel beta sheets. The strands of the  
CC beta sheets run parallel to the fiber axis. Long stretches of silk  
CC fibroin are composed of microcrystalline arrays of (-Gly-Ser-Gly-  
CC Ala-Gly-Ala-)n interrupted by regions containing bulkier residues.  
CC The fiber is composed of microcrystalline arrays alternating with  
CC amorphous regions.  
CC -I- PTM: The interchain disulfide bridge is essential for the  
CC intracellular transport and secretion of fibroin.  
-----  
CC This Swiss-Protein entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
-----  
CC EMBL AF226688; AAF76983.1; -; Genomic DNA.  
DR EMBL V00094; CA234433.1; -; Genomic DNA.  
DR EMBL V00097; CA234433.1; -; Genomic DNA.  
DR EMBL S74439; AAB31861.1; -; mRNA.  
DR EMBL X13869; CA232076.1; -; mRNA.  
DR EMBL M5378; AAA22839.1; -; mRNA.  
DR EMBL AB017369; BAA33147.1; -; Genomic DNA.  
DR EMBL CK538369; -; NOT\_ANNOTATED\_CDS; mRNA.  
DR EMBL AADK01000575; -; NOT\_ANNOTATED\_CDS; Genomic DNA.  
DR PIR; S01844; S01844.  
KM Repeat; Signal; Silk.  
FT SIGNAL 1 21 Potential.  
FT CHAIN 22 5263 Fibroin heavy chain.  
FT REGION 149 5206 Highly repetitive.  
FT DISULFID 5244 5244 Interchain (with C-190 in light chain).  
FT DISULFID 5260 5263























[illegible]

DR	GO: 0006817; P:phosphate transport; IEA.
DR	InterPro; IPR008160; Collagen.
FT	NON_TER 1
SQ	SEQUENCE 907 AA; 72392 MW; A0864ECT08740A00 CRC64;
Query Match	29.5%; Score 1230.5; DB 2; Length 907;
Best Local Similarity	43.2%; Pred. No. 1.8e-66;
Matches 368; Conservative	12; Mismatches 334; Indels 137; Gaps 50;

[illegible]

AC O8C918;  
 DT 01-MAR-2003 (TReMBLrel. 23, Created)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
 DT 10-MAY-2005 (TReMBLrel. 30, Last annotation update)  
 DE Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched  
 DE library, clone:A630042119 product:elastin, full insert sequence  
 DE (Elaesin).  
 GN Name=ElN;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Murinae; Mus.  
 OC NCBI\_Taxid=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadoya K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
 RA Knehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schirml L. M., Staudli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barth G.,  
 RA Blake J., Boiffelli D., Bojunga N., Carninci P., de Bonaldo M. F.,  
 RA Brownstein M. J., Bult C., Fletcher C., Fujita M., Gojobori T.,  
 RA Guncionich S., Hill D. P., Bult C., Hume D. A., Kamita M., Lee N. H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monnaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K. H., Weltz C., Whitaker C., Wilming L.,  
 RA Wyszynski B. S., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaide I., Otsu N., Saito R., Suzuki H., Yamana K. I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D. P., Bult C., Hume D. A., Quackenbush J.,  
 RA Schirml L. M., Kanapin A., Matsuda H., Batalov S., Beisel K. W.,  
 RA Blake J. A., Bradt D., Brusic V., Chothia C., Corbani L. E., Cousins S.,  
 RA Dalla E., Dragani T. A., Fletcher C. F., Forrest A., Frazer K. S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough U.,  
 RA Grimmond S., Guncionich S., Hirokawa N., Jackson I. J., Jarvis E. D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R. M., King B. L.,  
 RA Kono A. A., Kurochkin I. V., Lee Y., Lenhard B., Lyons P. A.,  
 RA Maglott D. R., Malais L., Marchionni L., McKenzie L. J., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W. J., Pettes G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J. O., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J. C., Reed D. J., Reid J., Ring B. Z., Ringwald M.,  
 RA Sadelin A., Schneider C., Sempole C. A., Setou M., Shimada K.,  
 RA Sultana R., Takekura Y., Taylor M. S., Teasdale R. D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L. G., Wyszynski B. S., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E. S., Rogers J.,

RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 50,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,  
 RA Kono H., Akiyama J., Nishi K., Kitanai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R. L., Collin F. S., Wagner L., Shenman C. M., Schuler G. D.,  
 RA Klausner R. D., Zeeberg B., Buetow K. H., Schaefer C. F., Bhat N. K.,  
 RA Altschul S. F., Jordan H., Moore T., Wax S. J., Wang J., Heien F.,  
 RA Hopkins R. F., Jordan B., Buetow K. H., Schaefer C. F., Bhat N. K.,  
 RA Diatchenko L., Marusina K., Farmer A. A., Rubin G. M., Hong L.,  
 RA Scapleton M., Soares M. B., Bonaldo M. F., Casavant T. L., Scheetz T. E.,  
 RA Brownstein M. J., Urdin T. B., Tohyuki S., Carninci P., Prange C.,  
 RA Raba S. S., Loggellano N. A., Peters G. J., Abramson R. D., Mullaly S. J.,  
 RA Bosak S. A., McEwan P. J., McKernan K. J., Malek J. A., Gunaratne P. H.,  
 RA Richards S., Worley K. C., Hale S., Garcia A. M., Gay L. V., Huiy S. W.,  
 RA Villalón D. K., Muzny D. M., Sodergren E. J., Lu X., Gibbs R. A.,  
 RA Fahney J., Halton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A. C., Shevchenko Y., Bouffard G. G.,  
 RA Blakesley R. W., Touchman J. W., Green E. D., Dickson M. C.,  
 RA Rodriguez A. C., Grimwood J., Schmutz J., Myers R. M.,  
 RA Butterfield Y. S. N., Krzywnicki M. I., Skalska U., Smalins D. E.,  
 RA Schnerch A., Schrein J. E., Jones S. J. M., Marra M. A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;  
 RX Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AK041860; BAC31084.1; -; mRNA.  
 DR EMBL; BC051649; AA051649.1; -; mRNA.

DR Ensemble: ENSEMBL0000029675; Mus musculus.  
DR MGI:95317; Eln.  
DR GO:GO:0007519; P:myogenesis; IMP.  
DR GO:GO:0030833; P:regulation of actin filament polymerization; IMP.  
DR GO:GO:0043149; P:stress fiber formation; IMP.  
DR InterPro:IPR003979; tropoelastin.  
DR PRINTS:PRO1500; TROPOLASTIN.  
SQ SEQUENCE 860 AA; 71938 MW; 7C340F2FFDC92E5 CRC64;

Query Match 29.2%; Score 1219.5; DB 2; Length 860;  
Best Local Similarity 44.1%; Pred. No. 7,7e-66;  
Matches 375; Conservative 50; Mismatches 295; Indels 131; Gaps 57;

QY 1 GAGAGSAGAGSGVG-----VPGVGVGV-GVPGKGVPGVPGVPGVGP-GVPGAGA 52  
DB 50 GAGTGGGAGGAGLGGGKPKPGAGLGTFTGAGPGGLGAGPGAGLGAFTPGAG 109  
QY 53 ---GSGAGAGSAGAGSAGAG--SGVGVPGVGVPGVPGKGVPGVPGVPGVPGV 108  
DB 110 LVPGGAAGAAAYKAAKAGAGLGVG---GVPG---GVGVGVPGVGVGVPG--- 158  
QY 109 GAGAGSAGAGSAGAGSAGAGSAGVGP--GVGVPGVPGKGVPGV--PGVGP--GVGP 164  
DB 159 GVGVGVPGVGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 217  
QY 165 GVG---PGAGAGSAGAGSAGAGSAGAGSAGVGPVG-----VPGV--GVGP----- 207  
DB 218 GTGARFPFVGVLPGVPTGTGTGKAKPGGGAGFAGTIPGVPGGQGPVPLGYPIKAPKL 277  
QY 208 -GKGVPGVPGVPGVPGVPGVPGAGAGSAGAGSAGAGSAGAGSAGVPGVPGV 266  
DB 278 GGYGLPTYNKLLPYGA--GAGKAGVPTGTGVSQAAAAAKAKAGAGAGV-LPGVG- 334  
QY 267 PGKGVPGVPGVPGVPGVPGVPGAGAGSAGAGSAGAGSAG--GAGSGV--GVPGVGP 322  
DB 335 -GGGIPG-----GAGAIPIGIGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 388  
QY 323 GVGVPG--KGVPGV--PGV--GPVPGAGAGSAGAGSAGAGSAGAGSAGVGP 377  
DB 389 GAGTIPGVGGTIPGVGGTIPGVGGTIPGVGGTIPGVGGTIPGVGGTIPGVGGT 448  
QY 378 GVGVPGVPGKGVPGVPGVPGVPGVPGVPGAGAGSAGAGSAGAGSAGVGP 437  
DB 449 TYGV-GAG-----GPGVGVAGAGLGG--GASPAATAAAKAKKAGAG---GAGALGGL--- 496  
QY 438 GVGVPGVPGKGVPGVPGVPGVPGVPGV--PGAGAGSAGAGSAGAGSAGV- 494  
DB 497 -----VPG-AVPGALPGAVPAV--FGAGGVPGAGTPAATAAAKAKAKAGLGPVG 546  
QY 495 GVGVPVPVPGVGP--GKGVPGVPGVPGV--VPGVPGV--AGAGSAGAGSAGAGS--- 546  
DB 547 GVPG-GVGVGGLPGVGVGVPGVPGVPGVGTGTGAGPGGLGAGSPAAAKSAKAKAAK 605  
QY 547 -GAGAGSAGVPGVPGVGP-----GKGVPGVPGV--PGVPGV--PGAGAGS--GAG 597  
DB 606 YRAAAGLGAAGVPGFGA--GAGVPGFAGAGVPGFAGAGVPGFAGAGVPGFAGAGV 664  
QY 598 AGSGA---GAGSAGAGSAGVGP--GVGVPGVPGKGVPGVPGVPGVPGVPGAG 653  
DB 665 AASKAAKYGAAGLGGPGGLGGPGGLGGPG--GLGAGVPGFVAGAAAP---PAAAAA 720  
QY 654 SGAGAGSAGAGSAGAGSAGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGAG 713  
DB 721 AARAAYGLGAGAGLGG--GLGAGGLGAGGLGAGGLGAG--GLGAG---GLGAG 773  
QY 714 SGAGAGSAGAGSAGAGSAGVPGV--GVPGVGP---PGKVPGVPGVPGVPGVGP 767  
DB 774 GGVSPAAAAKAKKAGAGLGG--GVLAGPFPGGVAAAPFGGLSPITPGGAG--GLGVGK 831  
QY 768 PGAGAGSAGAG 778  
DB 832 PPKPYGALGA 842

Search completed: May 10, 2006, 10:43:43  
Job time : 241 secs

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

## OM protein - protein search, using sw model

Run on: May 10, 2006, 10:36:19 ; Search time 189 Seconds

(Without alignments)  
1813.309 Million cell updates/sec

Title: US-10-800-179-19

Sequence: 1 GAAGSGAGAGSGVGPVG.....GVGPVGPGAGAGSGAGAGS 780

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapekt 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_21.\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*  
9: geneseqp2005s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4173	100.0	780	ADK51951	ADK51951 Repeat pr
2	4173	100.0	780	ADR70468	ADR70468 Silk-el
3	4173	100.0	780	ADU98730	ADU98730 Active ag
4	4173	100.0	780	ADU81661	ADU81661 Silk-el
5	4173	100.0	780	ABE57376	ABE57376 SEIP 47K
6	3448	82.6	884	AAR80341	AAR80341 Protein p
7	3448	82.6	884	AAW09213	AAW09213 SEIP8K po
8	3448	82.6	884	AAW53541	AAW53541 Expected
9	3448	82.6	884	AAW49728	AAW49728 SEIP8K po
10	3448	82.6	884	AAW51882	AAW51882 Plasmid p
11	3448	82.6	884	ABG31412	ABG31412 SEIP8K po
12	3448	82.6	884	ABW01628	ABW01628 Plasmid p
13	3448	82.6	884	ADR70473	ADR70473 Silk-el
14	3448	82.6	884	ADU98735	ADU98735 Active ag
15	3448	82.6	884	ADU81666	ADU81666 Silk-el
16	3448	82.6	884	ABE57381	ABE57381 SEIP 47E-
17	3448	82.6	983	ADU81678	ADU81678 Cotton bi
18	3448	82.6	1027	ADU81673	ADU81673 Cecropin
19	3448	82.6	1105	ADU81674	ADU81674 Antimicro
20	3448	82.6	1125	ADU81675	ADU81675 Green flu
21	3346	80.2	965	ADU81676	ADU81676 Active ag
22	3346	80.2	965	ADU81677	ADU81677 Active ag
23	3346	80.2	965	ADU81678	ADU81678 Active ag
24	3346	80.2	965	ABE57387	ABE57387 SEIP 58 r

25	3305	79.2	1038	8	ADR70478	ADR70478 Silk-el
26	3305	79.2	1038	9	ADU98740	ADU98740 Active ag
27	3305	79.2	1038	9	ADU81671	ADU81671 Silk-el
28	3305	79.2	1038	9	ABE57386	ABE57386 SEIP 67K
29	3296	79.0	2257	1	AAE82961	AAE82961 SEIP3 pro
30	3296	79.0	2257	2	AAE41012	AAE41012 SEIP3 mul
31	3296	79.0	2257	2	AAW26347	AAW26347 SEIP3 syn
32	3296	79.0	2257	2	AAW53523	AAW53523 Amino aci
33	3296	79.0	2257	3	AAE78282	AAE78282 SEIP3 ami
34	3296	79.0	2257	5	ABG69272	ABG69272 SEIP3/Elas
35	3296	79.0	2257	7	ABE44977	ABE44977 Recombina
36	3284	78.7	877	2	AAE80335	AAE80335 Protein p
37	3284	78.7	877	2	AAW49724	AAW49724 Protein p
38	3264	78.2	766	5	ABE53466	ABE53466 SEIP8 rel
39	3263	78.2	1016	9	ADU81677	ADU81677 Radioprot
40	3263	78.2	1016	9	ABE57388	ABE57388 SEIP47K-P
41	3253	78.0	832	5	ABE80252	ABE80252 Polymer S
42	3253	78.0	832	5	ABE53473	ABE53473 Protein p
43	3250	77.9	1043	9	ADU81676	ADU81676 Cotton bi
44	3192	76.5	768	5	ABE53481	ABE53481 Protein p
45	3182	76.3	988	2	AAE80253	AAE80253 Polymer S

## ALIGNMENTS

RESULT 1  
ADK51951  
ID ADK51951 standard; protein; 780 AA.  
XX  
AC ADK51951;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Repeat protein polymer repeat sequence, SEQ ID 19.  
XX  
KM repeat protein.  
XX  
OS Unidentified.  
XX  
PN MO2003099465-A1.  
XX  
PD 04-DEC-2003.  
XX  
PF 20-MAY-2003; 2003MO-US015757.  
XX  
PR 20-MAY-2002; 2002US-0381913P.  
XX  
PA (DOMO ) DOM CORNING CORP.  
XX (GENV ) GENENCOR INT INC.  
XX  
XX Kumar M;  
XX WPI; 2004-023263/02.  
XX  
XX Forming inorganic structure, for use as e.g. electronic photonic and nano  
XX composite materials, comprises providing substrate with repeat protein  
XX polymer, and exposing substrate to precursor containing inorganic  
XX species.  
XX  
XX Claim 12; Page 8; 27pp; English.  
XX  
XX The present invention relates to a method for forming an inorganic  
XX structure. The method comprises providing substrate with a repeat protein  
XX polymer, and exposing the substrate to a precursor comprising inorganic  
XX species. The repeat protein polymer catalyzes the formation of an  
XX inorganic structure on the substrate. The repeat portion of the repeat  
XX protein polymer may be ADK51951-ADK51957, which may have a head sequence  
XX at the N-terminal end (ADK51958) and a tail sequence at the C-terminal  
XX end. (ADK51959-ADK51966).  
SQ Sequence 780 AA;

	Query Match	100.0%;	Score 4173;	DB 8;	Length 780;	
	Best Local Similarity	100.0%;	Pred. No. 4,4e-264;			
	Matches 780;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
OY	1 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	60				
Db	1 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	60				
OY	61 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	120				
Db	61 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	120				
OY	121 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	180				
Db	121 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	180				
OY	181 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	240				
Db	181 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	240				
OY	241 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	300				
Db	241 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	300				
OY	301 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	360				
Db	301 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	360				
OY	361 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	420				
Db	361 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	420				
OY	421 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	480				
Db	421 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	480				
OY	481 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	540				
Db	481 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	540				
OY	541 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	600				
Db	541 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	600				
OY	601 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	660				
Db	601 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	660				
OY	661 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	720				
Db	661 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	720				
OY	721 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	780				
Db	721 GAGAGSGAAGAGSGVGVPGVGPVGKGVPCVGVBPVGVPBGVPGAAGAGSGAGAGS	780				

RESULT 2  
 ADR70468  
 ID ADR70468 standard; protein; 780 AA.  
 XX ADR70468;  
 AC  
 XX  
 DT 02-DEC-2004 (first entry)  
 DE Silk-elastin polymer SELP47K.  
 KW personal care composition; repeat sequence protein polymer;  
 KW hair care composition; shampoo; conditioner; anti-dandruff treatment;  
 KW styling aid; styling conditioner; hair repair treatment;  
 KW chemical hair treatment; hair dye; skin care composition; body wash;  
 KW antimicrobial cleanser; skin protectant cream; body lotion; facial cream;  
 KW moisturizing cream; facial cleansing emulsion;

KN	surfactant-based facial cleanser;facial exfoliating gel;
KM	anti-acne treatment; facial toner; exfoliating cream;facial mask;
KM	after shave balm; sunscreen; anti-fungal treatment; skin protectant;
KM	antiperspirant; cosmetic composition; makeup; eye gel;
KM	high-melting point lipstick; lipstick; lip gloss; lip balm; mascara;
KM	eyeliner; pressed powder formulation; foundation; nail care composition;
KM	nail enamel; cuticle treatment; nail polish; nail treatment;
KM	polish remover; oral care composition; toothpaste; mouth rinse;
KM	breath freshener; teeth whitening treatment;
KM	over-the-counter pharmaceutical composition; transparent film formation;
KM	hydrogel formation; water-solubility; luster; softness;
KM	moisture retinment; tensile property; viscoelastic behaviour;
KM	glass transition temperature; cloud temperature;
KM	decomposition temperature; silk; elastin; repeating sequence unit.
OS	Synthetic.
XX	
PN	US2004180027-A1.
PD	
XX	16-SEP-2004.
XX	
PF	12-MAR-2004; 2004US-00800179.
XX	
PR	12-MAR-2003; 2003US-0454077P.
XX	
PA	(KUMA/) KUMAR M.
PA	(CUEV/) CUEVAS W A.
XX	
PI	Kumar M, Cuevas WA;
XX	
DR	WPI; 2004-675584/66.
XX	
PT	Personal care composition useful as cosmetic, hair care or skin care
PT	product, comprises repeat sequence protein polymer and compounds such as
PT	carriers, excipients, liposomes, active ingredients, or emollients.
XX	
PS	Example 1, SEQ ID NO 19; 50pp; English.
XX	
CC	The invention describes a personal care composition (I) comprising a
CC	repeat sequence protein polymer with the balance of the composition
CC	comprising one or more compounds chosen from carriers, excipients,
CC	liposomes, active ingredients, biological or botanical products,
CC	humectants, emollients, surfactants, thickening agents, silicone
CC	components, organic sunscreens, preservatives, neutralising agents,
CC	perfumes and pigments. (I) is useful as a hair care composition such as
CC	shampoo, conditioner, anti-dandruff treatment, styling aids, styling
CC	conditioner, hair repair or treatment, serum, lotion, cream, pomade, or
CC	chemical treatment such as hair dye, skin care composition such as
CC	moisturising body wash, body wash, antimicrobial cleanser, skin
CC	protectant cream, body lotion, facial cream, moisturising cream, facial
CC	cleansing emulsion, surfactant-based facial cleanser, facial exfoliating
CC	gel, anti-acne treatment, facial toner, exfoliating cream, facial mask,
CC	after shave balm or sunscreen, skin care composition topically applied
CC	over-the-counter drugs comprising anti-fungal treatments, anti-acne
CC	treatments, skin protectants, and antiperspirants; cosmetic composition
CC	comprising a makeup composition chosen from eye gel, high-melting point
CC	lipstick, lipstick, lip gloss, lip balm, mascara, eyeliner, pressed
CC	powder formulation and foundation; nail care composition such as nail
CC	enamel, cuticle treatment, nail polish, nail treatment, or polish remover
CC	in an oral care composition such as toothpaste, mouth rinse, breath
CC	freshener, or whitening treatment; and over-the-counter pharmaceutical
CC	composition. The cosmetic composition is a mascara such as non-waterproof
CC	mascara, waterproof mascara, volumising mascara, lengthening mascara,
CC	curling mascara, anhydrous waterproof mascara, water-based mascara, or
CC	eyelash or eyebrow treatment; a pressed powder formulation such as loose
CC	powder, blush, eye shadow, or bronzing powder; foundation such as water-
CC	in-oil foundation, water-in-silicone foundation, oil-in-water foundation,
CC	anhydrous makeup stick, or cream-to-powder foundation. (I) has desired
CC	characteristics such as transparent film formation, hydrogel formation,
CC	better efficacy and binding to skin, hair, nail, and oral surfaces,
CC	desired level of hydrophobicity with water-solubility, imparting luster,
CC	softness, moisture retinment, and mechanical properties (such as tensile
CC	properties, viscoelastic behaviour glass transition temperature, cloud



QY	421	GAGAGSGAGAGSGVGPVGVPBGVPGKGVGVBGGVGPVCVCPGVGPAGAAGSGAGAGS	480
Dd	421	GAGAGSGAGAGSGVGPVGVPBGVPGKGVGBGVGGVGPVGGVPPGAAGSGAGAGS	480
QY	481	GAGAGSGAGAGSGVGPVGVPBGVPGKGVGBGVGGVGPVGGVPPGAAGSGAGAGS	540
Dd	481	GAGAGSGAGAGSGVGPVGVPBGVPGKGVGBGVGGVGPVGGVPPGAAGSGAGAGS	540
QY	541	GAGAGSGAGAGSGVGPVGVPBGVPGKGVGBGVGGVGPVGGVPPGAAGSGAGAGS	600
Dd	541	GAGAGSGAGAGSGVGPVGVPBGVPGKGVGBGVGGVGPVGGVPPGAAGSGAGAGS	600
QY	601	GAGAGSGAGAGSGVGPVGVPBGVPGKGVGBGVGGVGPVGGVPPGAAGSGAGAGS	660
Dd	601	GAGAGSGAGAGSGVGPVGVPBGVPGKGVGBGVGGVGPVGGVPPGAAGSGAGAGS	660
QY	661	GAGAGSGAGAGSGVGPVGVPBGVPGKGVGBGVGGVGPVGGVPPGAAGSGAGAGS	720
Dd	661	GAGAGSGAGAGSGVGPVGVPBGVPGKGVGBGVGGVGPVGGVPPGAAGSGAGAGS	720
QY	721	GAGAGSGAGAGSGVGPVGVPBGVPGKGVGBGVGGVGPVGGVPPGAAGSGAGAGS	780
Dd	721	GAGAGSGAGAGSGVGPVGVPBGVPGKGVGBGVGGVGPVGGVPPGAAGSGAGAGS	780

## RESULT 4

ID ADU81661 standard; protein; 780 AA.

AC ADU81661;

DT 10-FEB-2005 (first entry)

DE Silk-elastin polymer SLP47K, SEQ ID NO:19.

**KW** Protein engineering; protein stabilization; cosmetics; textile; filter;

KW fibroin; elastin.

OS Bombyx mori

OS Chimeric.

XX

XX

XX

XX

XX

PA (CUEV/) CUEVAS W A.

XX

XX

✕

PT composition, cosmetic, oral care composition, comprising conjugation

XX

XX

CC sequence protein polymer (i.e., an engineered protein) and at least

CC sunscreens, lanolin, vitamins, bleaches, thickening agents,

CC protein polymer preferably comprises a repeating amino acid sequence unit

CC dragline silk, gluten high molecular subunit, titin, fibronectin,

CC	lentinulinlidiin, glue polypeptide, ice nucleating protein, keratin,
CC	mucin, RNA polymerase II, resilin or mixtures thereof. The repeat
CC	sequence protein polymers provide a scaffolding to which active agents
CC	may be bound. In particular, they enhance the ability to utilize proteins
CC	and peptides as active agents as they overcome many of the
CC	hydrophobicity, solubility, and binding limitations associated with
CC	polypeptide active agents, and allow the use of polypeptide active agents
CC	in applications for which a total protein environment is desirable. The
CC	invention also relates to personal care compositions comprising a
CC	biomolecular conjugate of the invention, a method of making the personal
CC	care compositions; a method of making a biomolecular conjugate by
CC	recombinantly producing a fusion protein comprising the repeat sequence
CC	protein polymer and a polypeptide active agent; biomaterials adapted for
CC	at least one predetermined desirable function which comprise one or more
CC	biomolecular conjugates comprising a silk elastin polymer as the repeat
CC	sequence protein polymer and a polypeptide active agent; and a method of
CC	producing such biomaterials. The biomolecular conjugates are useful in
CC	personal care compositions such as over-the-counter pharmaceutical
CC	compositions, and compositions used for hair care, skin care, nail care,
CC	oral care and cosmetics. The biomolecular conjugates are also useful in
CC	biomaterials such as textiles (e.g., protective garments), filters,
CC	creams, coatings, and dressings, and in biomaterials useful in genetics
CC	research tools or in search and/or identification tools. The present
CC	sequence represents silk-elastin polymer SESP47K, a repeat sequence
CC	protein polymer comprising silk fibroin and elastin repeat sequences
CC	which is specifically claimed for use in biomolecular conjugates of the
CC	invention.
XX	
SO	Sequence 780 AA;
Query Match	100.0%; Score 4173; DB 9; Length 780;
Best Local Similarity	100.0%; Prid. No. 4,4e-264;
Matches 780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	

**SQ** Sequence 780 AA;

Query Match	100.0%;	Score 4173;	DB 9;	Length 780;
-------------	---------	-------------	-------	-------------

```
Matches 780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

[illegible]



[illegible]



Dd		394	GVCVPGVCPGAPAGAGSGAGAGSAGAGAGSGAGAGSGCVCVPGVGVPGVGPVGKGVPG	455
Oy		398	VG-PGVG-PGVG-PGAGAGSGAGAGSGAGAGSGAGSGGVCVPGVGVPGKGVPGV	454
Dd		454	VGVPGVGPVGVPBGAGAGSGAGAGSGAGAGSGAGSGGVCVPGVGVPGVGPVK	513
Oy		455	G-PGVG-PGVG-PGVG-PGAGAGSGAGAGSGAGAGSGAGSGGVCVPGVGPVK	510
Dd		514	GVPGVGPVGVPVGVPVPGAGAGSGAGAGSGAGAGSGAGSGGVCVPGVGPVK	573
Oy		511	VPGVG-PGVG-PGVG-PGVG-PGAGAGSGAGAGSGAGAGSGAGSGGVCVPGVGP	566
Dd		574	VPGVGPVGVPVGVPVGVPVPGAGAGSGAGAGSGAGAGSGAGSGGVCVPGVGP	633
Oy		567	PGKVPVG-PGVG-PGVG-PGVG-PGAGAGSGAGAGSGAGAGSGAGSGGVCVP	622
Dd		634	PQVGPVGKGVPGVGPVGVPVPGAGAGSGAGAGSGAGAGSGAGSGGVCVP	693
Oy		623	GVGVPGKGVPGV-GPVG-PGVG-PGVG-PGAGAGSGAGAGSGAGAGSGAGSGVPG	678
Dd		694	GVGVPGVGPVGKGVPGVGPVGVPVPGAGAGSGAGAGSGAGAGSGGVCVP	753
Oy		679	VGVGPVGVPVGKGVPGV-GPVG-PGVG-PGVG-PGAGAGSGAGAGSGAGAGSGV	734
Dd		754	VGVGPVGVPVGVPVGKGVPGVGPVGVPVPGAGAGSGAGAGSGAGAGSGV	813
Oy		735	GVPGVGPVGVPVGKGVPGV-GPVG-PGVG-PGVG-PGAGAGSGAGAGS	780
Dd		814	GVPGVGPVGVPVGVPVGVPVGKGVPGVGPVGVPVPGAGAGSGAGAGS	863
RESULT 8				
ID	AAM53541	AAM53541 standard; protein; 884 AA.		
AC	AAM53541;			
DT	10-AUG-1998	(first entry)		
DE	Expected amino acid sequence of pPT0345 encoding SELP8K polymer.			
XX	XX	KM	Polymer SELP8K; peptide repeat unit; DNA repeat unit;	
XX	XX	KW	high molecular weight polymer; synthetic silk; silk worm.	
OS	Synthetic.			
XX	XX	WO9810063-A1.		
FN	PD	12-MAR-1998.		
XX	PF	23-SEP-1996;	96MO-US015306.	
XX	PR	03-SEP-1996;	96US-00707237.	
PA	(PROT-)	PROTEIN POLYMER TECHNOLOGIES INC.		
PI	Ferrari FA,	Cappello J,	Crissman JM,	Dortman MA;
DR	WPI; 1998-193613/17.			
XX				
PT	Preparation of synthetic repetitive DNA - useful for construction of			
PT	large protein polymers having repeating units, used in structural			
PT	material, e.g. synthetic silk.			
XX				
PS	Example 7; Page 89; 127pp; English.			
XX				
CC	This is the amino acid sequence of pPT0345 comprising the SELP8K protein,			
CC	used in the method of invention, which involves the preparation of			
CC	synthetic DNA sequence having repeating units from about 3-15 codons and			
CC	encoding a protein of at least about 30 kDa. The method is useful for the			
CC	production of high molecular weight polymers (e.g. synthetic silk),			
CC	either nucleic acids or peptides that are the expression products of the			
CC	nucleic acids and particularly high molecular weight peptides containing			



	aa having a functional group reactive with at least one of aldehyde, iso(thio) cyanate and activated carboxy. (I) have similar biocompatibility to fibrin glues, but set more quickly and give a bond with greater shear strength. They are made from readily available natural sources, are easy to administer and are gradually resorbed. This sequence represents a crosslinking protein polymer unit designated SELP8K
CC	aa having a functional group reactive with at least one of aldehyde,
CC	iso(thio) cyanate and activated carboxy. (I) have similar biocompatibility
CC	to fibrin glues, but set more quickly and give a bond with greater shear
CC	strength. They are made from readily available natural sources, are easy
CC	to administer and are gradually resorbed. This sequence represents a
CC	crosslinking protein polymer unit designated SELP8K
XX	
SQ	Sequence 884 AA;
Query Match	82.6%; Score 3448; DB 3; Length 884;
Best Local Similarity	90.6%; Pred. No. 6, 9e-217;
Matches	752; Conservative 0; Mismatches 26; Indels 52; Gaps 52;
QY	3 GAGSGAGAGSGGVPGVPGVPGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGSGAGA 58
DB	34 GAGSGAGAGSGGVPGVPGVPGVPGVPGVGKGVPGVPGVPGVPGAGAGSGAGA 93
QY	59 GSGAGAGSGAGAGSGGVPGVPGVPGVPGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGS 114
DB	94 GSGAGAGSGAGAGSGGVPGVPGVPGVPGVPGVPGKGVPGVPGVPGVPGAGAGS 153
QY	115 GAGAGSGAGAGSGAGAGSGGVPGVPGVPGKGVPGVG-PGVG-PGVG-PGVG-PGA 170
DB	154 GAGAGSGAGAGSGAGAGSGGVPGVPGVPGVPGVPGKGVPGVPGVPGVPGAG 213
QY	171 GAGSGAGAGSGAGAGSGAGSGGVPGVPGVPGKGVPGVG-PGVG-PGVG-PGVG 227
DB	214 GAGSGAGAGSGAGAGSGAGSGGVPGVPGVPGVPGVPGKGVPGVPGVPGVG 273
QY	228 -PGAAGSGAGAGSGAGAGSGAGSGGVPGVPGVPGVPGKGVPGVG-PGVG-PGVG- 283
DB	274 VPGAAGSGAGAGSGAGAGSGAGSGGVPGVPGVPGVPGVPGVPGKGVPGVPGVPGV 333

[illegible]

XX	ABG31412;
AC	
XX	
DT	29-NOV-2002 (first entry)
XX	
DE	SELP8K polymer encoded by plasmid pPRO345.
XX	
KW	Protein polymer; functional group; crosslink; sealing; filling; tissue;
KW	tissue mass; tissue bonding; resorbable bond; flexible bond; sealant;
KW	adhesive; wound healing; burn dressing; blood flow; ruptured vessel;
KW	artery; vein; structural protein; vunerary; fibroin; elastin; collagen;
KW	keratin; SELP8K.
XX	
OS	Unidentified.
OS	Synthetic.
XX	
PN	US6423333-B1.
XX	
XX	23-JUL-2002.
PD	
XX	
PF	29-NOV-1999; 99US-00451206.
XX	
PR	05-MAY-1995; 95US-00435641.
PR	02-MAY-1996; 96US-00642246.
XX	
PA	(PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
XX	
PI	Stedronsky ER, Cappello J;
XX	
DR	WPI; 2002-672937/72.
XX	
PT	Crosslinked protein composition used as sealant or adhesive for sealing
PT	or filling defect in viable tissue, as burn dressing, or in wound healing
PT	e.g. to staunch flow of fluid e.g. blood, through ruptured vessels.
XX	
PS	Example 2; Col 17; 46pp; English.
XX	
CC	The present invention relates to protein polymers having repetitive units
CC	from naturally occurring structural proteins such as fibroin, elastin,
CC	collagen and keratin. The polymers comprise a functional group which can
CC	be chemically crosslinked with appropriate crosslinkers. The protein
CC	polymer is produced by recombinant DNA technology. The protein polymer is
CC	useful for sealing or filling a defect in viable tissue, particularly for
CC	augmenting tissue mass. The protein polymer is useful in a variety of
CC	applications related to their physical, chemical and biological
CC	properties, and/or to bond together separated tissue to provide a stable,
CC	flexible or resorbable bond. The protein is particularly useful as a
CC	sealant or adhesive, in wound healing or as a burn dressing e.g. to stop
CC	or staunch the flow of fluid (e.g. blood) through ruptured vessels (e.g.
CC	arteries or veins). The present sequence represents SELP8K polymer
XX	
XX	
XX	Sequence 884 AA;
XX	

Query Match	82.6%;	Score 3448;	DB 5;	Length 884;
Best Local Similarity	90.6%;	Pred. No. 6.9e-217;		
Matches 752; Conservative	0;	Mismatches 26;	Indels 52;	Gaps 52

[illegible][illegible]

```

RESULT 12
ABW01628
ID ABW01628 standard; protein; 884 AA.
XX
AC ABW01628;
XX
DT 12-FEB-2004 (first entry)
XX
DE Plasmid pPRO345 SELPAK polymer protein.
XX
KW Crosslinked protein; sealant; depot; drug; therapy; gel; film; thread;
KM coating; vulneryary.
XX
OS Unidentified.
XX
PN US2003104589-A1.
XX
PD 05-JUN-2003.
XX
PF 05-APR-2002; 2002US-00117931.
XX
PR 05-MAY-1995; 95US-00435641.
XX PR 02-MAY-1996; 96US-00642246.
XX PR 29-NOV-1999; 99US-00451206.
XX
PA (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
XX
PI Stedronsky ER, Cappello J;
XX
DR WPI; 2003-829350/77.
XX
PT Crosslinked protein composition useful as sealing a defect in tissue, the

```

PT protein prior to crosslinking comprises repetitive units of 3-15 amino  
PT acids of natural structural protein.

XX Example 2; Page 27-29; Opp; English.

XX The present invention relates to crosslinked protein composition. The  
CC invention is useful as sealants or depots to provide for relatively  
CC uniform release of a physiologically active product e.g., drug and for  
CC the formation of articles of manufacture such as gels, films, threads,  
CC coatings. The present sequence is plasmid pPTO345 SEP8K polymer protein  
XX  
XX Sequence 884 AA;

Query Match 82.6%; Score 3448; DB 7; Length 884;  
Best Local Similarity 90.6%; Pred. No. 6.9e-217;  
Matches 752; Conservative 0; Mismatches 26; Indels 52; Gaps 52;

QY 3 GAGSGAGSGVPGVGVPGVPGKGVPGVG-PGVG-PGVG-PGVG-PGAGSGAGA 58  
DB 34 GAGSGAGSGVPGVGVPGVPGVPGKGVPGVPGVPGVPGVPGVPGVPGVPGV 93  
QY 59 GSGAGSGAGSGSGVPGVPGVPGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGS 114  
DB 94 GSGAGSGAGSGSGVPGVPGVPGKGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 153  
QY 115 GAGAGSGAGSGAGSGSGVPGVPGVPGKGVPGVG-PGVG-PGVG-PGVG-PGA 170  
DB 154 GAGAGSGAGSGAGSGSGVPGVPGVPGKGVPGVPGVPGVPGVPGVPGVPGV 213  
QY 171 GAGAGSGAGSGAGSGAGSGVPGVPGVPGKGVPGVG-PGVG-PGVG-PGVG 227  
DB 214 GAGSGAGSGAGSGAGSGAGSGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 273  
QY 228 -PGAGSGAGSGAGSGAGSGAGSGVPGVPGVPGKGVPGVG-PGVG-PGVG- 283  
DB 274 VPGAGAGSGAGSGAGSGAGSGAGSGVPGVPGVPGVPGVPGVPGVPGVPGV 333  
QY 284 PGVGV-PGAGAGSGAGSGAGSGAGSGAGSGVPGVPGVPGVPGVPGVPGV 340  
DB 334 PGVGVPGAGAGSGAGSGAGSGAGSGAGSGVPGVPGVPGVPGVPGVPGV 393  
QY 341 GVG-PGVG-PGAGAGSGAGSGAGSGAGSGAGSGVPGVPGVPGVPGVPGV 397  
DB 394 GVGPGVGVPGAGAGSGAGSGAGSGAGSGAGSGVPGVPGVPGVPGVPGV 453  
QY 398 VG-PGVG-PGVG-PGAGAGSGAGSGAGSGAGSGAGSGVPGVPGVPGVPGV 454  
DB 454 VGVPGVGVPGVPGVPGAGAGSGAGSGAGSGAGSGAGSGVPGVPGVPGV 513  
QY 455 G-PGVG-PGVG-PGAGAGSGAGSGAGSGAGSGAGSGVPGVPGVPGVPGV 510  
DB 514 GVPVGVPGVPGVPGVPGVPGAGAGSGAGSGAGSGAGSGAGSGVPGVPGV 573  
QY 511 VPGVG-PGVG-PGVG-PGAGAGSGAGSGAGSGAGSGAGSGVPGVPGVPGV 566  
DB 574 VPGAGVPGVPGVPGVPGVPGAGAGSGAGSGAGSGAGSGAGSGVPGVPGV 633  
QY 567 PGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGSGAGSGAGSGAGSGAGSGV 622  
DB 634 PGVGVPGKGVPGVPGVPGVPGVPGAGAGSGAGSGAGSGAGSGAGSGV 693  
QY 623 GVGVPKGVPGVG-PGVG-PGVG-PGVG-PGAGAGSGAGSGAGSGAGSGV 678  
DB 694 GVGVPKGVPGVPGVPGVPGVPGVPGAGAGSGAGSGAGSGAGSGAGSGV 753  
QY 679 VGVGVGVPGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGSGAGSGAGSGAG 734  
DB 754 VGVGVGVPGVPGVPGKGVPGVPGVPGVPGVPGAGAGSGAGSGAGSGAG 813  
QY 735 GVPVGVPGVPGVPGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGSGAGSGAG 780  
DB 814 GVPVGVPGVPGVPGVPGKGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 863

RESULT 13

ID ADR70473 standard; protein; 884 AA.

AC ADR70473;

DT 02-DEC-2004 (first entry)

DE Silk-elastin polymer SBLP 47-E13.

KM personal care composition; repeat sequence protein polymer;  
KM hair care composition; shampoo; conditioner; anti-dandruff treatment;  
KM styling aid; styling conditioner; hair repair treatment;  
KM chemical hair treatment; hair dye; skin care composition; body wash;  
KM antimicrobial cleanser; skin protectant cream; body lotion; facial cream;  
KM moisturizing cream; facial cleanser; facial exfoliating gel;  
KM surfactant-based facial cleanser; facial exfoliating gel;  
KM anti-acne treatment; facial toner; exfoliating cream; facial mask;  
KM after shave balm; sunscreen; anti-fungal treatment; skin protectant;  
KM antiperspirant; cosmetic composition; makeup; eye gel;  
KM high-melting point lipstick; lipstick; lip gloss; lip balm; mascara;  
KM eyeliner; pressed powder formulation; foundation; nail care composition;  
KM nail enamel; cuticle treatment; nail polish; nail treatment;  
KM polish remover; oral care composition; toothpaste; mouth rinse;  
KM breath freshener; teeth whitening treatment;  
KM over-the-counter pharmaceutical composition; transparent film formation;  
KM hydrogel formation; water-solubility; luster; softness;  
KM glass retinament; tensile property; viscoelastic behaviour;  
KM glass transition temperature; cloud temperature;  
KM decomposition temperature; silk; elastin; repeating sequence unit.

OS Synthetic.

US2004180027-A1.

PD 16-SEP-2004.

PN 12-MAR-2004; 2004US-00800179.

PR 12-MAR-2003; 2003US-0454077P.

PA (KUMA/) KUMAR M.

PA (CUEV/) CUEVAS W A.

PI Kumar M, Cuevas WA;

DR WPI; 2004-675584/66.

PT Personal care composition useful as cosmetic, hair care or skin care  
PT product, comprises repeat sequence protein polymer and compounds such as  
PT carriers, excipients, liposomes, active ingredients, or emollients.

PS Disclosure; SEQ ID NO 25; 50pp; English.

XX The invention describes a personal care composition (I) comprising a  
CC repeat sequence protein polymer with the balance of the composition  
CC comprising one or more compounds chosen from carriers, excipients,  
CC liposomes, active ingredients, biological or botanical products,  
CC humectants, emollients, surfactants, thickening agents, silicone  
CC components, organic sunscreens, preservatives, neutralizing agents,  
CC perfumes and pigments. (I) is useful as a hair care composition such as  
CC shampoo, conditioner, anti-dandruff treatment, styling aids, styling  
CC conditioner, hair repair or treatment, serum, lotion, cream, pomade, or  
CC chemical treatment such as hair dye; skin care composition such as  
CC moisturizing body wash, body wash, antimicrobial cleanser, skin  
CC protectant cream, body lotion, facial cream, moisturizing cream, facial  
CC cleansing emulsion, surfactant-based facial cleanser, facial exfoliating  
CC gel, anti-acne treatment, facial toner, exfoliating cream, facial mask,  
CC after shave balm or sunscreen; skin care composition topically applied  
CC over-the-counter drugs comprising anti-fungal treatments, anti-acne  
CC treatments, skin protectants, and antiperspirants; cosmetic composition  
CC comprising a makeup composition chosen from eye gel, high-melting point



CC lipstick, lipstick, lip gloss, lip balm, mascara, eyeliner, pressed  
CC powder formulation and foundation; nail care composition such as nail  
CC enamel, cuticle treatment, nail polish, nail treatment, or polish remover  
CC , an oral care composition, such as toothpaste, mouth rinse, breath  
CC freshener, or whitening treatment; and over-the-counter pharmaceutical  
CC composition. The cosmetic composition is a mascara such as non-waterproof  
CC mascara, waterproof mascara, volumizing mascara, lengthening mascara,  
CC curling mascara, anhydrous waterproof mascara, water-based mascara, or  
CC eyelash or eyebrow treatment; a pressed powder formulation such as loose  
CC powder, blush, eye shadow, or bronzing powder; foundation such as water-  
CC in-oil foundation, water-in-silicone foundation, oil-in-water foundation,  
CC anhydrous makeup stick, or cream-to-powder foundation. (I) has desired  
CC characteristics such as transparent film formation, hydrogel formation,  
CC better efficacy and binding to skin, hair, nail, and oral surfaces,  
CC desired level of hydrophobicity with water-solubility, imparting luster,  
CC softness, moisture retentment, and mechanical properties (such as tensile  
CC properties, viscoelastic behaviour, glass transition temperature, cloud  
CC temperature and decomposition temperature), and does not have any  
CC chemical modifications of the protein. This is the amino acid sequence of  
CC silk-elastin polymer SLP 47-E13 that may be used as the repeat sequence  
CC protein polymer of the invention.

XX Sequence 884 AA;

SQ Query Match 82.6%; Score 3448; DB 8; Length 884;

Best Local Similarity 90.6%; Pred. No. 6.9e-217;

Matches 752; Conservative 0; Mismatches 26; Indels 52; Gaps 52;

QY 3 GAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - PGVG - PGAGAGSAGA 58  
DB 34 GAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - PGVG - PGAGAGSAGA 93  
QY 59 GSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - PGVG - PGAGAGS 114  
DB 94 GSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - PGVG - PGAGAGS 153  
QY 115 GAGAGSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - PGVG - PGA 170  
DB 154 GAGAGSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - PGVG - PGA 213  
QY 171 GAGSGAGAGSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - PGVG - PGVG 227  
DB 214 GAGSGAGAGSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - PGVG - PGVG 273  
QY 228 - PGAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - PGVG - 283  
DB 274 VPAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - PGVG - 333  
QY 284 PGVG - PGAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - P 340  
DB 334 PGVGVPAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - P 393  
QY 341 GVG - PGVG - PGAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG - PG 397  
DB 394 GVGEPGVPAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - 453  
QY 398 VG - PGVG - PGAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGV 454  
DB 454 VGVPAGVPAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - 513  
QY 455 G - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPVGVPKGVP 510  
DB 514 GVGVPAGVPAGVPAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG 573  
QY 511 VPGVG - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPVGVP 566  
DB 574 VPBEPGVPAGVPAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG 633  
QY 567 PGVGVPVG - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPVG 622  
DB 634 PGVGVPBEPGVPAGVPAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPVGVPKGVP 693  
QY 623 GVGVPAGVPAGVP - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGAGAGSGVGPVG 678

DB 694 GVGVPAGVPAGVPAGVPAGVPAGAGSGAGAGSGAGAGSGAGAGSGVGPVG 753  
QY 679 VGVPAGVPAGVPAGVPAGVPAGVPAGAGSGAGAGSGAGAGSGVGPVG 754  
DB 754 VGVPAGVPAGVPAGVPAGVPAGVPAGAGSGAGAGSGAGAGSGVGPVG 813  
QY 735 GVGVPAGVPAGVPAGVPAGVPAGVPAGVPAGAGSGAGAGSGVGPVG 780  
DB 814 GVGVPAGVPAGVPAGVPAGVPAGVPAGVPAGAGSGAGAGSGVGPVG 863

## RESULT 14

ID ADU98735 standard; protein; 884 AA.

AC ADU98735;

DT 10-FRB-2005 (first entry)

DE Active agent release control system-related protein SeqID25.

KM cosmetics; delivery mechanism.

OS Unidentified.

PN US2004228913-A1.

PD 18-NOV-2004.

PF 14-MAY-2004; 2004US-00845775.

PR 14-MAY-2003; 2003US-0470465P.

PA (KUMA/) KUMAR M.

PA (MAZE/) MAZEAUD I.

PI (CHR/) CHRISTIANO S P.

PI Kumar M, Mazeaud I, Christiano SP;

DR WPI; 2005-037015/04.

PT System for providing controlled release delivery of active agent useful

PT for incorporating active agents into personal care product compositions,

PT comprises repeat sequence protein polymer and active agent.

PS Disclosure; SEQ ID NO 25; 34pp; English.

XX This invention relates to a novel system for providing controlled release

XX delivery of an active agent which comprises a repeat sequence protein

XX polymer and at least one active agent. The invention may be useful in a

XX hair care composition, skin care composition, nail care composition,

XX cosmetic composition, oral care composition or over-the-counter

XX pharmaceutical composition. The system may be useful in shampoos, gels,

XX mousses and other hair care products, rinse-off conditioners, skin care

XX products such as moisturizers, toners and makeup and nail care products

XX such as polishes and polish removers. The invention enables controlled

XX release of active agents. The present sequence is that of a protein which

XX is related to the invention.

SQ Sequence 884 AA;

Query Match 82.6%; Score 3448; DB 9; Length 884;

Best Local Similarity 90.6%; Pred. No. 6.9e-217;

Matches 752; Conservative 0; Mismatches 26; Indels 52; Gaps 52;

QY 3 GAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - PGVG - PGAGAGSAGA 58  
DB 34 GAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - PGVG - PGAGAGSAGA 93  
QY 59 GSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - PGVG - PGAGAGS 114  
DB 94 GSGAGAGSGAGAGSGVGPVGVPVGVPKGVPVG - PGVG - PGVG - PGAGAGS 153





Db	214	GAGSGAGAGSGAGAGSGAGAGSGVGPVGVPGVGPVGVPGBGVPGVGPVGVPGB	273
QY	228	- PGAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPGVPGKGPVG- PGVG- PGVG-	283
Db	274	VPBAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPGVPGVGPGBGVPGVGPGB	333
QY	284	PGVG- PGAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPGVPGKGPVG- PGVG- P	340
Db	334	PGVGPBAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPGVPGVGPGBGVPGVGP	393
QY	341	GVG- PGVG- PGAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPGVPGKGPVG- PG	397
Db	394	GVGPBGPVGAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPGVPGVGPGBGVPG	453
QY	398	VG- PGVG- PGAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPGVPGKGPVG	454
Db	454	VGPBGPVGPGVGPBAGAGSGAGAGSGAGAGSGVGPVGVPGVPGVPGVPGV	513
QY	455	G- PGVG- PGVG- PGAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPGVPGKGP	510
Db	514	GVGPVGVPVGVPBGPVGPGAGAGSGAGAGSGAGAGSGVGPVGVPGVPGVPGV	573
QY	511	VPBVG- PGVG- PGVG- PGAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVPGV	566
Db	574	VPBGPVGVPVGVPBGPVGPGAGAGSGAGAGSGAGAGSGVGPVGVPGVPGV	633
QY	567	PGKGPVG- PGVG- PGVG- PGAGAGSGAGAGSGAGAGSGAGAGSGVGPVGVP	622
Db	634	PGVGPBGPVGPGVGPBGPVGPGAGAGSGAGAGSGAGAGSGVGPVGVP	693
QY	623	GVGPBGPVG- PGVG- PGVG- PGAGAGSGAGAGSGAGAGSGAGAGSGVGPVG	678
Db	694	GVGPBGPVGPGVGPBGPVGPGVGPBGPVGPGAGAGSGAGAGSGAGAGSGVPG	753
QY	679	VGPBGPVGPGKGPVG- PGVG- PGVG- PGAGAGSGAGAGSGAGAGSGV	734
Db	754	VGPBGPVGPGVGPBGPVGPGVGPBGPVGPGAGAGSGAGAGSGAGAGSGV	813
QY	735	GVGPVGVPVGVPBGPVG- PGVG- PGVG- PGAGAGSGAGAGSG	780
Db	814	GVGPVGVPVGVPBGPVGPGVGPBGPVGPGAGAGSGAGAGSG	863

Search completed: May 10, 2006, 10:39:41  
Job time : 195 secs

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2006, 10:43:59 ; Search time 47 Seconds  
(without alignments)  
1372.064 Million cell updates/sec

Title: US-10-800-179-19

Perfect score: 4173  
Sequence: 1 GAAGSGAGSGAGSGVGVPGVC.....GVGVGVPGAGSGAGSG 780

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

- Database : Issued Patents AA:\*
- 1: /cgn2\_6/ptodata/1/1aa/5\_COMB.pep:\*
  - 2: /cgn2\_6/ptodata/1/1aa/6\_COMB.pep:\*
  - 3: /cgn2\_6/ptodata/1/1aa/H\_COMB.pep:\*
  - 4: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*
  - 5: /cgn2\_6/ptodata/1/1aa/RE\_COMB.pep:\*
  - 6: /cgn2\_6/ptodata/1/1aa/backfilltest.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3526	84.5	889	2 US-08-806-029-19	Sequence 19, Appl
2	3448	82.6	884	1 US-08-397-633A-68	Sequence 68, Appl
3	3448	82.6	884	1 US-08-435-641-15	Sequence 15, Appl
4	3448	82.6	884	1 US-08-707-237A-96	Sequence 96, Appl
5	3448	82.6	884	2 US-08-642-246-15	Sequence 15, Appl
6	3448	82.6	884	2 US-09-451-206-15	Sequence 15, Appl
7	3448	82.6	884	4 PCT-US96-06229-15	Sequence 15, Appl
8	3296	79.0	2257	1 US-08-175-155-47	Sequence 47, Appl
9	3296	79.0	2257	1 US-08-477-509B-82	Sequence 82, Appl
10	3296	79.0	2257	1 US-08-707-237A-53	Sequence 82, Appl
11	3296	79.0	2257	2 US-08-482-085B-82	Sequence 82, Appl
12	3296	79.0	2257	2 US-09-444-791A-82	Sequence 82, Appl
13	3284	78.7	877	1 US-08-397-633A-54	Sequence 54, Appl
14	3253	78.0	832	1 US-08-212-237-4	Sequence 4, Appl
15	3253	78.0	832	2 US-08-806-029-27	Sequence 27, Appl
16	3253	78.0	832	4 PCT-US95-02772-4	Sequence 4, Appl
17	3192	76.5	768	2 US-08-806-029-35	Sequence 35, Appl
18	3182	76.3	988	1 US-08-212-237-5	Sequence 5, Appl
19	3182	76.3	988	2 US-08-806-029-28	Sequence 28, Appl
20	3182	76.3	988	4 PCT-US95-02772-5	Sequence 28, Appl
21	3176	76.1	1413	1 US-08-175-155-39	Sequence 39, Appl
22	3176	76.1	1413	1 US-08-707-237A-45	Sequence 45, Appl
23	3176	76.1	1464	1 US-08-477-509B-74	Sequence 74, Appl
24	3176	76.1	1464	2 US-08-482-085B-74	Sequence 74, Appl
25	3158	75.7	2055	1 US-09-444-791A-74	Sequence 74, Appl
26	3158	75.7	2055	1 US-08-175-155-46	Sequence 46, Appl
27	3158	75.7	2055	1 US-08-477-509B-81	Sequence 81, Appl

28	3158	75.7	2055	1 US-08-707-237A-52	Sequence 52, Appl
29	3158	75.7	2055	2 US-08-482-085B-81	Sequence 81, Appl
30	3158	75.7	2055	2 US-09-444-791A-80	Sequence 80, Appl
31	3089	74.0	2018	2 US-09-444-791A-81	Sequence 80, Appl
32	3089	74.0	2100	1 US-08-477-509B-80	Sequence 80, Appl
33	3089	74.0	2100	2 US-08-482-085B-80	Sequence 80, Appl
34	3089	74.0	2107	1 US-08-175-155-45	Sequence 45, Appl
35	3089	74.0	2107	1 US-08-707-237A-51	Sequence 51, Appl
36	2998	71.8	1056	1 US-08-212-237-6	Sequence 6, Appl
37	2998	71.8	1056	2 US-08-806-029-29	Sequence 29, Appl
38	2998	71.8	1056	4 PCT-US95-02772-6	Sequence 6, Appl
39	2855	68.4	1002	1 US-08-707-237A-103	Sequence 103, App
40	2855	68.4	1002	2 US-08-642-246-25	Sequence 25, Appl
41	2855	68.4	1002	2 US-09-451-206-25	Sequence 25, Appl
42	2855	68.4	1002	4 PCT-US96-06229-25	Sequence 25, Appl
43	2830	67.8	750	2 US-08-806-029-16	Sequence 25, Appl
44	2814	67.4	696	2 US-08-806-029-16	Sequence 36, Appl
45	2795	67.0	936	1 US-08-707-237A-108	Sequence 108, App

ALIGNMENTS

RESULT 1  
US-08-806-029-19  
Sequence 19, Application US/08806029  
Patent No. 6380154  
GENERAL INFORMATION:  
APPLICANT: Cappello, Joseph  
APPLICANT: Stedronsky, Erwin R.  
TITLE OF INVENTION: Synthetic Proteins for in vivo Drug  
TITLE OF INVENTION: Delivery and Tissue Augmentation  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hobbach, Teet, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: United States  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/806,029  
FILING DATE: 24-FEB-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/212,237  
FILING DATE: 11-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Tregarten, Richard F.  
REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-58847-2/RFT/MTK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 889 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-806-029-19  
Query Match 84.5%; Score 3526; DB 2; Length 889;  
Best Local Similarity 92.2%; Pred. No. 5.6e-254;  
Matches 765; Conservative 0; Mismatches 13; Indels 52; Gaps 52;  
Gy 3 GAAGSGAGSGAGSGVGVPGVC-PGVG-PGVG-PGVG-PGAGSGAGA 58



[illegible]

```

1      RESULT 3
2      US-08-435-641-15
3      ; Sequence 15, Application US/08435641
4      ; Patent No. 5817303
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Sledronsky, Erwin R.
7      ; TITLE OF INVENTION: Tissue Adhesive Using Synthetic
8      ; TITLE OF INVENTION: Crosslinking
9      ; NUMBER OF SEQUENCES: 15
10     ; CORRESPONDENCE ADDRESS:
11     ; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
12     ; STREET: Four Embarcadero Center, Suite 3400
13     ; CITY: San Francisco
14     ; STATE: CA
15     ; COUNTRY: US
16     ; ZIP: 94111
17     ; COMPUTER READABLE FORM:
18     ; MEDIUM TYPE: Floppy disk
19     ; COMPUTER: IBM PC compatible
20     ; OPERATING SYSTEM: PC-DOS/MS-DOS
21     ; SOFTWARE: Patent In Release #1.0, Version #1.30
22     ; CURRENT APPLICATION DATA:
23     ; APPLICATION NUMBER: US/08/435,641
24     ; FILING DATE: 05-MAY-1995
25     ; CLASSIFICATION: 514
26     ; ATTORNEY/AGENT INFORMATION:
27     ; NAME: Rowland, Berttram I.
28     ; REGISTRATION NUMBER: 20015
29     ; REFERENCE/DOCKET NUMBER: A-61127
30     ; TELECOMMUNICATION INFORMATION:
31     ; TELEPHONE: 415-781-1989
32     ; TELEFAX: 415-398-3249
33     ; INFORMATION FOR SEQ ID NO: 15:
34     ; SEQUENCE CHARACTERISTICS:
35     ; LENGTH: 884 amino acids
36     ; TYPE: amino acid
37     ; STRANDEDNESS: single
38     ; TOPOLOGY: linear
39     ; MOLECULE TYPE: peptide
40     ; JS-08-435-641-15

```

[illegible][illegible]

RESULT 4  
 US-08-707-237A-96  
 Sequence 96, Application US/08707237A  
 Patent No. 5830713  
 GENERAL INFORMATION:  
 APPLICANT: Ferrari, Franco A.  
 APPLICANT: Ferrari, Joseph  
 APPLICANT: Crisman, John W.  
 APPLICANT: Dorman, Mary A.  
 TITLE OF INVENTION: METHODS FOR PREPARING SYNTHETIC  
 TITLE OF INVENTION: REPETITIVE DNA  
 NUMBER OF SEQUENCES: 108  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Flehr, Hobach, Test, Albrighton & Herbert  
 STREET: Four Embarcadero Center, Suite 3400  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: United States  
 ZIP: 94111-4187  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent-in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/707,237A  
 FILING DATE: 03-SEP-1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/175,155  
 FILING DATE: 29-DEC-1993



QY 115 GAGAGSAGAGSAGAGSGVPGVPGVPGKGVPGVG-PGVG-PGVG-PGVG-PGA 170  
DB 154 GAGAGSAGAGSAGAGSGVPGVPGVPGVPGKGVPGVPGVPGVPGVPGV 213  
QY 171 GAGSAGAGSAGAGSAGAGSGVPGVPGVPGKGVPGVG-PGVG-PGVG-PGVG 227  
DB 214 GAGSAGAGSAGAGSAGAGSGVPGVPGVPGVPGKGVPGVPGVPGVPGV 273  
QY 228 -PGAAGSAGAGSAGAGSGVPGVPGVPGKGVPGVG-PGVG-PGVG- 283  
DB 274 VPAGAGSAGAGSAGAGSGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 333  
QY 284 PGVG-PGAGSAGAGSAGAGSGVPGVPGVPGVPGVPGVPGVPGV-PGVG-P 340  
DB 334 PGVPGAGAGSAGAGSAGAGSGVPGVPGVPGVPGVPGVPGVPGVPGV 393  
QY 341 GVG-PGAGSAGAGSAGAGSGVPGVPGVPGVPGVPGVPGVPGVPGV-PG 397  
DB 394 GVGPGVPGAGAGSAGAGSGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 453  
QY 398 VG-PGVG-PGAGSAGAGSAGAGSGVPGVPGVPGVPGVPGVPGVPGV 454  
DB 454 VGVPGVPGVPGAGAGSAGAGSGVPGVPGVPGVPGVPGVPGVPGVPGV 513  
QY 455 G-PGVG-PGVG-PGAGSAGAGSAGAGSGVPGVPGVPGVPGVPGVPGV 510  
DB 514 GVPVPGVPGVPGVPGVPGAGAGSAGAGSGVPGVPGVPGVPGVPGV 573  
QY 511 VPGVG-PGVG-PGVG-PGAGSAGAGSAGAGSGVPGVPGVPGVPGV 566  
DB 574 VPGAGVPGVPGVPGVPGVPGAGAGSAGAGSGVPGVPGVPGVPGV 633  
QY 567 PGKVPGVG-PGVG-PGVG-PGAGSAGAGSAGAGSGVPGVPGVPGV 622  
DB 634 PGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 693  
QY 623 GVGPGKGVPGV-PGVG-PGVG-PGVG-PGAGSAGAGSAGAGSGVPGV 678  
DB 694 GVGPGVPGKGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 753  
QY 679 VGVPGVPGKGVPGV-PGVG-PGVG-PGVG-PGAGSAGAGSAGAGSGV 734  
DB 754 VGVPGVPGVPGKGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 813  
QY 735 GVPVPGVPGVPGKGVPGV-PGVG-PGVG-PGVG-PGAGSAGAGS 780  
DB 814 GVPVPGVPGVPGVPGVPGKGVPGVPGVPGVPGVPGVPGVPGVPGV 863

RESULT 6  
US-09-451-206-15  
; Sequence 15, Application US/09451206  
; Patent No. 6423333  
; GENERAL INFORMATION:  
; APPLICANT: STEDRONSKY, Erwin R.  
; CAPPELLO, Joseph  
; TITLE OF INVENTION: Tissue Adhesive Using Synthetic  
; Crosslinking  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
; STREET: Four Embarcadero Center, Suite 200  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/451,206  
FILING DATE: 29-Nov-6423333-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/642,246  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: ROWLAND, Bertram I  
REGISTRATION NUMBER: 20015  
REFERENCE/DOCKET NUMBER: A61127-1/BIR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 884 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
US-09-451-206-15  
Query Match 82.6%; Score 3448; DB 2; Length 884;  
Best Local Similarity 90.6%; Pred. No. 3.3e-248;  
Matches 752; Conservative 0; Mismatches 26; Indels 52; Gaps 52;  
QY 3 GAGSAGAGSAGAGSGVPGVPGVPGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGSAGA 58  
DB 34 GAGSAGAGSAGAGSGVPGVPGVPGVPGKGVPGVPGVPGVPGVPGVPGVPGVPGV 93  
QY 59 GAGSAGAGSAGAGSGVPGVPGVPGVPGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGS 114  
DB 94 GAGSAGAGSAGAGSGVPGVPGVPGVPGVPGKGVPGVPGVPGVPGVPGVPGVPGV 153  
QY 115 GAGSAGAGSAGAGSGVPGVPGVPGVPGKGVPGVG-PGVG-PGVG-PGVG-PGA 170  
DB 154 GAGSAGAGSAGAGSGVPGVPGVPGVPGKGVPGVPGVPGVPGVPGVPGVPGVPGV 213  
QY 171 GAGSAGAGSAGAGSGVPGVPGVPGVPGKGVPGVG-PGVG-PGVG-PGVG-PGVG 227  
DB 214 GAGSAGAGSAGAGSGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 273  
QY 228 -PGAAGSAGAGSAGAGSGVPGVPGVPGVPGKGVPGVG-PGVG-PGVG- 283  
DB 274 VPAGAGSAGAGSAGAGSGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 333  
QY 284 PGVG-PGAGSAGAGSAGAGSGVPGVPGVPGVPGVPGVPGVPGV-PGVG-P 340  
DB 334 PGVPGAGAGSAGAGSAGAGSGVPGVPGVPGVPGVPGVPGVPGVPGV 393  
QY 341 GVG-PGAGSAGAGSAGAGSGVPGVPGVPGVPGVPGVPGVPGVPGV-PG 397  
DB 394 GVGPGVPGAGAGSAGAGSGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 453  
QY 398 VG-PGVG-PGAGSAGAGSAGAGSGVPGVPGVPGVPGVPGVPGVPGV 454  
DB 454 VGVPGVPGVPGAGAGSAGAGSGVPGVPGVPGVPGVPGVPGVPGVPGV 513  
QY 455 G-PGVG-PGVG-PGAGSAGAGSAGAGSGVPGVPGVPGVPGVPGVPGV 510  
DB 514 GVPVPGVPGVPGVPGVPGAGAGSAGAGSGVPGVPGVPGVPGVPGVPGV 573  
QY 511 VPGVG-PGVG-PGVG-PGAGSAGAGSAGAGSGVPGVPGVPGVPGVPGV 566  
DB 574 VPGAGVPGVPGVPGVPGVPGAGAGSAGAGSGVPGVPGVPGVPGVPGV 633  
QY 567 PGKVPGVG-PGVG-PGVG-PGAGSAGAGSAGAGSGVPGVPGVPGVPGV 622  
DB 634 PGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 693  
QY 623 GVGPGKGVPGV-PGVG-PGVG-PGVG-PGAGSAGAGSAGAGSGVPGV 678





[illegible]

```

;      TOPOLOGY: linear
;      MOLECULE TYPE: peptide

```



```

Db      1058 GAGGSGAGGSGAGGSGAGGSGAGGSGAGAAGSSGVGPVGPGV    1117
QY      745 GVPGKVPBGVG-PGVG-PGVG-PGAAGSAGAGS   760
       ||| |||| | |||| | |||| | |||| | |||| | |||| 
Db      1118 GVPGVPGVPGVPGVPGVPGVPGVPGVPGAAGSAGAGS   1157
RESULT 11
US-08-482-085B-82
; Sequence 82, Application US/08482085B
; Patent No. 6018030
GENERAL INFORMATION:
APPLICANT: Ferrari, Franco A.
APPLICANT: Richardson, Charles
APPLICANT: Chambers, James
APPLICANT: Causey, Stuart
APPLICANT: Pollock, Thomas J.
APPLICANT: Cappelletto, Joseph
APPLICANT: Crissman, John W.
TITLE OF INVENTION: No. 6018030e1 Peptides Comprising Repetitive
TITLE OF INVENTION: Units of Amino Acids and DNA Sequences Encoding the Same
NUMBER OF SEQUENCES: 112
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohnbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,085B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/927,258
FILING DATE: 04-NOV-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/114,618
FILING DATE: 29-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/053,049
FILING DATE: 22-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/175,155
FILING DATE: 29-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Treacartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE//DOCKET NUMBER: A-55186-6/RFT/MTX
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 2257 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-482-085B-82
Query Match          79.0%; Score 3296; DB 2; Length 2257;
Best Local Similarity 67.4%; Pred. No. 1.3e-236;
Matches 755; Conservative 0; Mismatches 25; Indels 340; Gaps 52;

```

[illegible]

```

RESULT 12
US-09-444-791A-82
; Sequence 82, Application US/09444791A
; Patent No. 6355776
GENERAL INFORMATION:
APPLICANT: Ferrari!, Franco A.
Richardson, Charles
Chambers, James
Cauley, Stuart
Pollock, Thomas J.
Cappelletto, Joseph
Crissman, John W.
TITLE OF INVENTION: No. 6355776el Peptides Comprising Repetitive
Units of Amino Acids and DNA Sequences Encoding the Same
NUMBER OF SEQUENCES: 117
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr Hohbach Test Albritton & Herbert LLP
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/444,791A
FILING DATE: 22-Nov. 6355776-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/482,085
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/175,155
FILING DATE: 29-DEC-1993
APPLICATION NUMBER: US 08/053,049
FILING DATE: 22-APR-1993
APPLICATION NUMBER: US 07/114,618
FILING DATE: 29-OCT-1987
APPLICATION NUMBER: US 06/927,258
FILING DATE: 04-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: Treacartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-55186-11/RFT/BTC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 2257 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 82:
US-09-444-791A-82

Query Match          79.0%; Score 3296; DB 2; Length 2257;
Best Local Similarity 67.4%; Pred. No. 1,3e-236;
Matches 755; Conservative 0; Mismatches 25; Indels 340; Gaps 52;

QY      1  GAGAGGAGAGGGVGVPGVPGVPGKGPVYG-PGVG-PGVG-PGV----- 46
           ||| | | | | | | | | | | | | | | | | | | | | | | |
DB       38  GAGAGSGAGAGGGVGVPGVPGVPGVPGVPGVPGVPGVPGVPGAGASGA   97
           ||| | | | | | | | | | | | | | | | | | | | | | | |
QY      47 -----GRPAGAGSAGAGSSGAGAGSGAGSGGAGSVGPVGVPKGAVP    92
           ||| | | | | | | | | | | | | | | | | | | | | | | |
DB       98  GAGSGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGGAGVPGVPGVGP    157

```



[illegible]

```

1      RESULT 14
2      US-08-212-237-4
3      Sequence 4, Application US/08212237
4      Patent No. 5606019
5      GENERAL INFORMATION:
6      APPLICANT: Cappello, Joseph
7      TITLE OF INVENTION: Synthetic Proteins As Implantables
8      NUMBER OF SEQUENCES: 9
9      CORRESPONDENCE ADDRESS:
10     ADDRESSSE: Flehr, Hohbach, Test, Albritton & Herbert
11     STREET: Four Embarcadero Center, Suite 3400
12     CITY: San Francisco
13     STATE: CA
14     COUNTRY: U.S.A.
15     ZIP: 94111-4187
16     COMPUTER READABLE FORM:
17     MEDIUM TYPE: Floppy disk
18     COMPUTER: IBM PC compatible
19     OPERATING SYSTEM: PC-DOS/MS-DOS
20     SOFTWARE: Patentin Release #1.0, Version #1.25
21     CURRENT APPLICATION DATA:
22     APPLICATION NUMBER: US/08/212.237
23     FILING DATE: 11-MAR-1994
24     CLASSIFICATION: 435
25     ATTORNEY/AGENT INFORMATION:
26     NAME: Rowland, Berttram I
27     REGISTRATION NUMBER: 20,015
28     REFERENCE/DOCKET NUMBER: A-58847/BIR
29     TELECOMMUNICATION INFORMATION:
30     TELEPHONE: 415-781-1989
31     TELEFAX: 415-398-3249
32     INFORMATION FOR SEQ ID NO: 4:
33     SEQUENCE CHARACTERISTICS:
34     LENGTH: 832 amino acids
35     TYPE: amino acid
36     STRANDEDNESS: single
37     TOPOLOGY: linear
38     MOLECULE TYPE: protein
39     US-08-212-237-4

```

	Query Match	78.0%	Score 3253	DB 1:	Length 832;
	Best Local Similarity	88.9%	Pred. No. 8-5e-234;		
	Matches 742; Conservative	0;	Mismatches 13;	Indels 80;	Gaps 65;
Oy	9 GAGSGVGVPGVGVPDGVGPKGVKGPVGVG-PGVG-PGVG-PGVGPGAGSAGAGGSACAG 65				
bcb	3 G-----VGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGV-GAGSAGSAGAGSAGAG 57				

[illegible]

```

RESULT 15
US-08-806-029-27
; Sequence 27, Application US/08806029
; Patent No. 6380154
;
; GENERAL INFORMATION:
;
; APPLICANT: Cappello, Joseph
; APPLICANT: Stedronsky, Erwin R.
; TITLE OF INVENTION: Synthetic Proteins for in vivo Drug
; TITLE OF INVENTION: Delivery and Tissue Augmentation
;
; NUMBER OF SEQUENCES: 36
;
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Flehr, Hohnbach, Test, Albrighton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
;
; ZIP: 94111
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;

```



**This Page Blank (uspto)**





D <b>b</b>	301	GAGAGSGAGAGSGVGP	PGVGP	PGVGP	PGKGP	CGVGP	BGPVGP	PGVGP	PGVGP	PGAGAGSGAGAGAS	360
Q <b>y</b>	361	GAGAGSGAGAGSGVGP	PGVGP	PGVGP	PGKGP	CGVGP	BGPVGP	PGVGP	PGVGP	PGAGAGSGAGAGAS	420
D <b>b</b>	361	GAGAGSGAGAGSGVGP	PGVGP	PGVGP	PGKGP	CGVGP	BGPVGP	PGVGP	PGVGP	PGAGAGSGAGAGAS	420
Q <b>y</b>	421	GAGAGSGAGAGSGVGP	PGVGP	PGVGP	PGKGP	CGVGP	BGPVGP	PGVGP	PGVGP	PGAGAGSGAGAGAS	480
D <b>b</b>	421	GAGAGSGAGAGSGVGP	PGVGP	PGVGP	PGKGP	CGVGP	BGPVGP	PGVGP	PGVGP	PGAGAGSGAGAGAS	480
Q <b>y</b>	481	GAGAGSGAGAGSGVGP	PGVGP	PGVGP	PGKGP	CGVGP	BGPVGP	PGVGP	PGVGP	PGAGAGSGAGAGAS	540
D <b>b</b>	481	GAGAGSGAGAGSGVGP	PGVGP	PGVGP	PGKGP	CGVGP	BGPVGP	PGVGP	PGVGP	PGAGAGSGAGAGAS	540
Q <b>y</b>	541	GAGAGSGAGAGSGVGP	PGVGP	PGVGP	PGKGP	CGVGP	BGPVGP	PGVGP	PGVGP	PGAGAGSGAGAGAS	600
D <b>b</b>	541	GAGAGSGAGAGSGVGP	PGVGP	PGVGP	PGKGP	CGVGP	BGPVGP	PGVGP	PGVGP	PGAGAGSGAGAGAS	600
Q <b>y</b>	601	GAGAGSGAGAGSGVGP	PGVGP	PGVGP	PGKGP	CGVGP	BGPVGP	PGVGP	PGVGP	PGAGAGSGAGAGAS	660
D <b>b</b>	601	GAGAGSGAGAGSGVGP	PGVGP	PGVGP	PGKGP	CGVGP	BGPVGP	PGVGP	PGVGP	PGAGAGSGAGAGAS	660
Q <b>y</b>	661	GAGAGSGAGAGSGVGP	PGVGP	PGVGP	PGKGP	CGVGP	BGPVGP	PGVGP	PGVGP	PGAGAGSGAGAGAS	720
D <b>b</b>	661	GAGAGSGAGAGSGVGP	PGVGP	PGVGP	PGKGP	CGVGP	BGPVGP	PGVGP	PGVGP	PGAGAGSGAGAGAS	720
Q <b>y</b>	721	GAGAGSGAGAGSGVGP	PGVGP	PGVGP	PGKGP	CGVGP	BGPVGP	PGVGP	PGVGP	PGAGAGSGAGAGAS	780
D <b>b</b>	721	GAGAGSGAGAGSGVGP	PGVGP	PGVGP	PGKGP	CGVGP	BGPVGP	PGVGP	PGVGP	PGAGAGSGAGAGAS	780

## RESULT 2

```

US-10-800-179-19
: Sequence 19, Application US/10800179
: Publication No. US20040180027A1
: GENERAL INFORMATION:
: APPLICANT: Cuevas, William A.
: APPLICANT: Kumar, Manoj
: TITLE OF INVENTION: Use of Repeat Sequence Protein Polymers in Personal Care
: TITLE OF INVENTION: Compositions
: FILE REFERENCE: DOC 0057 PA / GC792-4 / DC 5074
: CURRENT APPLICATION NUMBER: US/10/800.179
: CURRENT FILING DATE: 2004-03-12
: PRIOR APPLICATION NUMBER: 60/454,077
: PRIOR FILING DATE: 2003-03-12
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 19
: LENGTH: 780
: TYPE: PRT
: ORGANISM: Unknown
: FEATURE:
: OTHER INFORMATION: SELL 47K
: US-10-800-179-19

```

Query Match	100.0%;	Score 4173;	DB 4;	Length 780;
Best Local Similarity	100.0%;	Pred. No. 6.6e-248;		
Matches 780; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	1	60
QY	GAGAGSSGAGAGS	60
Dp	1	60
QY	61	120
Dp	61	120
QY	121	180
Dp	121	180
QY	181	240

Db	181	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	240
Qy	241	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	300
Db	241	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	300
Qy	301	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	360
Db	301	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	360
Qy	361	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	420
Db	361	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	420
Qy	421	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	480
Db	421	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	480
Qy	481	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	540
Db	481	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	540
Qy	541	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	600
Db	541	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	600
Qy	601	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	660
Db	601	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	660
Qy	661	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	720
Db	661	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	720
Qy	721	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	780
Db	721	GAGAGS	GAGAGS	GVGV	PGV	PGV	PGK	PGV	PGV	PGV	PGV	PGAGAGS	GAGAGS	780

### RESULT 3

```

US-10-845-775A-19
/ Sequence 19, Application US/10845775A
/ Publication No. US20040228913A1
/ GENERAL INFORMATION:
/   APPLICANT: Kumar, Manoj
/   APPLICANT: Mazeaud, Isabelle
/   APPLICANT: Christiano, Steven P.
/   TITLE OF INVENTION: Controlled Release of Active Agents Utilizing Repeat Sequence
/   TITLE OF INVENTION: Protein Polymers
/   FILE REFERENCE: DOC 0077 PA/GC 792-6/DC 5110
/   CURRENT APPLICATION NUMBER: US/10/845,775A
/   CURRENT FILING DATE: 2004-05-14
/   PRIOR APPLICATION NUMBER: US 60/470,465
/   PRIOR FILING DATE: 2003-05-14
/   NUMBER OF SEQ ID NOS: 31
/   SOFTWARE: PatentIn version 3.2
/   SEQ ID NO 19
/   LENGTH: 780
/   TYPE: PRT
/   ORGANISM: unknown
/   FEATURE:
/   OTHER INFORMATION: silk and elastin peptide repeats
US-10-845-775A-19

```

Query Match	100.0%	Score 4173	DB 5	Length 780
Best Local Similarity	100.0%	Pred. No. 6.6e-248		
Matches 780	0	Mismatches	0	Indels 0
Conservative	0	Gaps	0	

Oy	1	GAGAGSGAGAGSGVGVPGVGPVGIVPVGKGVPGVGPVGIVPVGVPVGPAGAGSGAGAGS	60
Dd	1	GAGAGSGAGAGSGVGVPGVGPVGIVPVGKGVPGVGPVGIVPVGVPVGPAGAGSGAGAGS	60
Oy	61	GAGAGSGAGAGSGVGVPGVGPVGIVPVGKGVPGVGPVGIVPVGVPVGPAGAGSGAGAGS	120





	Best Local Similarity	90.6%	Pred. No. 1.5e-203;	
	Matches	752;	Conservative	0; Mismatches 26; Indels 52; Gaps 52;
QY	3	GAGSAGAGSGSVGPVGVPVGKGVPGVG - PGVG - PGVG - PGVG - PGAGAGSGAGA	58	
Db	34	GAGSAGAGSGSVGPVGVPVGKGVPGVG - PGVG - PGVG - PGVG - PGAGAGSGAGA	93	
QY	59	GSAGAGSGAGAGSGVGPVGVPVGKGVPGVG - PGVG - PGVG - PGVG - PGAGAGS	114	
Db	94	GSAGAGSGAGAGSGVGPVGVPVGKGVPGVG - PGVG - PGVG - PGVG - PGAGAGS	153	
QY	115	GAGAGSGAGAGSGAGSGVGPVGVPVGKGVPGVG - PGVG - PGVG - PGVG - PGA	170	
Db	154	GAGAGSGAGAGSGAGSGVGPVGVPVGKGVPGVG - PGVG - PGVG - PGVG - PGA	213	
QY	171	GAGSAGAGSGAGAGSGAGSGVGPVGVPVGKGVPGVG - PGVG - PGVG - PGVG	227	
Db	214	GAGSAGAGSGAGAGSGAGSGVGPVGVPVGKGVPGVG - PGVG - PGVG - PGVG	273	
QY	228	- PGAGAGSGAGAGSGAGAGSGVGPVGVPVGKGVPGVG - PGVG - PGVG -	283	
Db	274	VPAGAGSGAGAGSGAGAGSGVGPVGVPVGKGVPGVG - PGVG - PGVG -	333	
QY	284	PGVG - PGAGAGSGAGAGSGAGAGSGVGPVGVPVGKGVPGVG - PGVG - P	340	
Db	334	PGVG - PGAGAGSGAGAGSGAGAGSGVGPVGVPVGKGVPGVG - PGVG - P	393	
QY	341	GVG - PGVG - PGAGAGSGAGAGSGAGAGSGVGPVGVPVGKGVPGVG - PG	397	
Db	394	GVG - PGVG - PGAGAGSGAGAGSGAGAGSGVGPVGVPVGKGVPGVG - PG	453	
QY	398	VG - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGVGPVGVPVGKGVPGV	454	
Db	454	VG - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGVGPVGVPVGKGVPGV	513	
QY	455	G - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGVGPVGVPVGKGVPGV	510	
Db	514	G - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGVGPVGVPVGKGVPGV	573	
QY	511	VPVG - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGVGPVGVPVGKGV	566	
Db	574	VPVG - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGVGPVGVPVGKGV	633	
QY	567	PGKGVPGVG - PGVG - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGVGPVG	622	
Db	634	PGKGVPGVG - PGVG - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGVGPVG	693	
QY	623	GVGPVGKGVPGVG - PGVG - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGVGPVG	678	
Db	694	GVGPVGKGVPGVG - PGVG - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGVGPVG	753	
QY	679	GVGPVGKGVPGVG - PGVG - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGVGPVG	734	
Db	754	GVGPVGKGVPGVG - PGVG - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGVGPVG	813	
QY	735	GVGPVGKGVPGVG - PGVG - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGVGPVG	780	
Db	814	GVGPVGKGVPGVG - PGVG - PGVG - PGVG - PGAGAGSGAGAGSGAGAGSGVGPVG	863	
RESULT 8				
US-10-800-179-25				
Sequence 25, Application US/10800179				
Publication No. US20040180027A1				
GENERAL INFORMATION:				
APPLICANT: Cuevas, William A.				
TITLE OF INVENTION: Use of Repeat Sequence Protein Polymers in Personal Care				
FILE REFERENCE: DOC 0057 PA / GC792-4 / DC 5074				
CURRENT FILING DATE: 2004-03-12				
PRIOR APPLICATION NUMBER: 60/454, 077				



[illegible]

OY		455 G-PGVG- PGVG- PGVG- P-GAGAGSAGAAGSGAAGSGTGVGPVGPVGVPKGG	510
Dd		514 GVPGVGPVGVEPVGVPGAAGAGSAGAAGSGAGSGTGVPVGVPVGPVG	573
OY		VPGVG- PGVG- PGVG- PGVG- P-GAGAGSAGAAGSGAAGSGTGVGPVGVPVG	566
Dd		574 VFGACVPBGVPPGVGEPEGVFPGAAGASGAAAGSGAAGSGTGVPCVTGPVG	633
OY		567 PKGVPVG- PGVG- PGVG- PGVG- PGAAGSGAAGSGAAGSGTAGSGTGVPGVGP	622
Dd		634 PGVGPBEGVPGVGPVGEPGVGPFGAAGSGAAGSGAAGSGTGVPGVGP	693
OY		623 GVGPVGKGVPGVG- PGYG- PGYG- PGYG- P-GAGAGSAGAAGSGAAGSGTGVPRG	678
Dd		694 GVGPVGVPBEBGPVGVPVGVEBPVGVPGAAGSGAAGSGAAGSGTGVPG	753
OY		679 VGVPGVGPVGKGVPGVG- PGVG- PGVG- PGVG- P-GAGAGSAGAAGSGAAGSGTGSV	734
Dd		754 VGVPGVGPVGVPBEGVPVGVPVGVBGPVGVPGAAGSGAAGSGAAGSGTGSV	813
OY		735 GVPGVGPVGVPVGKGVPGVG- PGVG- PGVG- PGVG- PGAAGSGAAGSG 780	
Dd		814 GVPGVGPVGVPVGVPBEGVPVGVPVGVEBPVGVPGAAGSGAAGSG 863	
RESULT 11			
		US-10-939-036-25	
	/ Sequence 25, Application US/10939036		
	/ Publication No. US20050142094A1		
	/ GENERAL INFORMATION:		
	/ APPLICANT: Kumar, Manoj		
	/ TITLE OF INVENTION: Use of Repeat Sequence Protein Polymers		
	/ FILE REFERENCE: GC792-4-2		
	/ CURRENT APPLICATION NUMBER: US/10/939,036		
	/ CURRENT FILING DATE: 2004-09-10		
	/ PRIOR APPLICATION NUMBER: US 10/800,179		
	/ PRIOR FILING DATE: 2004-03-12		
	/ PRIOR APPLICATION NUMBER: US 60/454,077		
	/ PRIOR FILING DATE: 2003-03-12		
	/ NUMBER OF SEQ ID NOS: 37		
	/ SOFTWARE: FastSBQ for Windows Version 4.0		
	/ SEQ ID NO 25		
	/ LENGTH: 884		
	/ TYPE: PRF		
	/ ORGANISM: Unknown		
	/ FEATURE:		
	/ OTHER INFORMATION: SLP 47E-13		
	US-10-939-036-25		
Query Match                  82.6%; Score 3448; DB 5; Length 884;			
	Best Local Similarity    90.6%; Pred. No. 1.5e-203;		
	Matches     752; Conservative    0; Mismatches    26; Indels      52; Gaps       52		
OY		3 GAGSGAGAGSGVGPVGVPVGKVPVG- PGVG- PGVG- PGVG- P-GAGAGSGAGA	58
Dd		34 GAGSGAGAGSGVGPVGVPVGVPVGVPBEGVPVGVPBEGVPVGVPGAAGSGAGA	93
OY		59 GSAGAGSGAGSGVGPVGVPVGKVPVG- PGYG- PGYG- PGYG- P-GAGAGS	114
Dd		94 GSGAGAGSGAGSGVGPVGVPBEGVPVGVPBEGVPVGVPBEGVPBEGVPBEGA	153
OY		115 GAGAGSGAGAGSGAGSGVPPGVGPVGVPVGKVPVG- PGVG- PGVG- PGVG- P-GA	170
Dd		154 GAGAGSGAGAGSGAGSGVPPGVGPVGVPBEGVPVGVPBEGVPVGVPBEGVP	213
OY		171 GAGSGAGAGSGAGAGSGAGSGVGPVGVPVGKVPVG- PGVG- PGVG- PGVG-	227
Dd		214 GAGSGAGAGSGAGAGSGAGSGVGPVGVPVGVPBEGVPVGVPBEGVPVG	273
OY		228 -PGAAGSGAAGSGAAGSGAGSGTGVPTGVTPKGVPVG- PGYG- PGYG- PGYG-	283
Dd		274 VPGAAGSGAAGSGAAGSGAGSGTGVPTGVTPKGVPVGVPBEGVPBEGVPBGE	333











GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2006, 10:45:38 ; Search time 113 Seconds

(without alignments)  
324.072 Million cell updates/sec

Title: US-10-800-179-19

Perfect score: 4173

Sequence: 1 GAGAGSGAGAGGAGVGPVG.....GVGPVGAGAGSGAGAGS 780

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 250354 seqs, 4694837 residues

Total number of hits satisfying chosen parameters: 250354

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:\*  
1: /SID55/ptodata/2/pubpaa/US06\_NEW\_PUB.pep1.\*  
2: /SID55/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
3: /SID55/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
4: /SID55/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
5: /SID55/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
6: /SID55/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
7: /SID55/ptodata/2/pubpaa/US09\_NEW\_PUB.pep1.\*  
8: /SID55/ptodata/2/pubpaa/US10\_NEW\_PUB.pep1.\*  
9: /SID55/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
10: /SID55/ptodata/2/pubpaa/US11\_NEW\_PUB.pep1.\*  
11: /SID55/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
12: /SID55/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3448	82.6	884	11 US-11-201-606-15	Sequence 15, Appl
2	2855	68.4	1002	11 US-11-201-606-25	Sequence 25, Appl
3	2795	67.0	936	11 US-11-201-606-30	Sequence 30, Appl
4	2740	65.7	966	11 US-11-201-606-34	Sequence 34, Appl
5	2091.5	50.1	1186	11 US-11-053-100-46	Sequence 46, Appl
6	2087.5	50.0	1309	11 US-11-053-100-53	Sequence 53, Appl
7	2086	50.0	939	11 US-11-053-100-26	Sequence 26, Appl
8	2086	50.0	1174	11 US-11-053-100-43	Sequence 43, Appl
9	2086	50.0	1225	11 US-11-053-100-50	Sequence 50, Appl
10	2086	50.0	1329	11 US-11-053-100-56	Sequence 56, Appl
11	1913.5	45.9	657	11 US-11-053-100-32	Sequence 32, Appl
12	1910.5	45.8	639	11 US-11-053-100-25	Sequence 25, Appl
13	1910.5	45.8	654	11 US-11-053-100-29	Sequence 29, Appl
14	1906.5	45.7	656	11 US-11-053-100-38	Sequence 38, Appl
15	1906.5	45.7	657	11 US-11-053-100-35	Sequence 35, Appl
16	1904	45.6	874	11 US-11-053-100-42	Sequence 42, Appl
17	1429.5	34.3	669	11 US-11-053-100-39	Sequence 39, Appl
18	1429.5	34.3	724	11 US-11-053-100-41	Sequence 41, Appl
19	1429.5	34.3	879	11 US-11-053-100-55	Sequence 55, Appl
20	1428.5	34.2	507	11 US-11-053-100-31	Sequence 31, Appl
21	1426	34.2	735	11 US-11-053-100-44	Sequence 44, Appl

22	1426	34.2	736	11 US-11-053-100-45	Sequence 45, Appl
23	1426	34.2	757	11 US-11-053-100-47	Sequence 47, Appl
24	1426	34.2	774	11 US-11-053-100-49	Sequence 49, Appl
25	1426	34.2	864	11 US-11-053-100-58	Sequence 58, Appl
26	1425	34.1	859	11 US-11-053-100-52	Sequence 52, Appl
27	1417.5	34.0	489	11 US-11-053-100-24	Sequence 24, Appl
28	1417.5	34.0	504	11 US-11-053-100-28	Sequence 28, Appl
29	1415	33.9	506	11 US-11-053-100-37	Sequence 37, Appl
30	1415	33.9	507	11 US-11-053-100-34	Sequence 34, Appl
31	1411	33.8	775	11 US-11-053-100-51	Sequence 51, Appl
32	1411	33.8	879	11 US-11-053-100-57	Sequence 57, Appl
33	1367.5	32.8	450	11 US-11-053-100-16	Sequence 16, Appl
34	1310	31.4	378	11 US-11-201-606-26	Sequence 26, Appl
35	1119.5	26.8	574	11 US-11-053-100-40	Sequence 40, Appl
36	1118.5	26.8	357	11 US-11-053-100-30	Sequence 30, Appl
37	1107.5	26.5	339	11 US-11-053-100-23	Sequence 23, Appl
38	1107.5	26.5	354	11 US-11-053-100-27	Sequence 27, Appl
39	1105	26.5	356	11 US-11-053-100-36	Sequence 36, Appl
40	1105	26.5	357	11 US-11-053-100-33	Sequence 33, Appl
41	1096	26.3	1901	11 US-11-052-554A-135	Sequence 135, App
42	1089.5	26.1	1079	11 US-11-052-554A-145	Sequence 145, App
43	1034	24.8	1660	11 US-11-052-554A-137	Sequence 137, App
44	1021	24.5	624	11 US-11-053-100-48	Sequence 48, Appl
45	1021	24.5	728	11 US-11-053-100-54	Sequence 54, Appl

#### ALIGNMENTS

RESULT 1  
US-11-201-606-15  
Sequence 15, Application US/11201606  
Publication NO. US20060029638A1  
GENERAL INFORMATION:  
APPLICANT: STEDRONSKY, Erwin R.  
TITLE OF INVENTION: Tissue Adhesive Using Synthetic  
TITLE OF INVENTION: Crosslinking  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
STREET: Four Embarcadero Center, Suite 200  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/11/201,606  
FILING DATE: 2005-08-10  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/642,246  
FILING DATE: 1996-05-02  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: ROWLAND, Bertram I  
REGISTRATION NUMBER: 20015  
REFERENCE/DOCKET NUMBER: A61127-1/BIR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 884 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-11-201-606-15

Query Match	82.6%;	Score 3448;	DB 11;	Length 884;
Best Local Similarity	90.6%;	Pred. No. 6.6e-230;		
Matches 752;	Conservative 0;	Mismatches 26;	Indels 52;	Gaps 52

QY	3	GAGSAGAGSGG	GGVGVGV	PGVGV	PGKGV	PGVG	-PGVG	-PGVG	-PGVG	-PGAAGS	GAGA	58
Db	34	GAGSAGAGSGG	GGVGV	PGVGV	PGKGV	PGVG	-PGVG	-PGVG	-PGVG	-PGAAGS	GAGA	93
QY	59	GSAGAGSGAGS	GGVGV	PGVGV	PGKGV	PGVG	-PGVG	-PGVG	-PGVG	-PGAAGS		114
Db	94	GSAGAGSGAGS	GGVGV	PGVGV	PGKGV	PGVG	-PGVG	-PGVG	-PGVG	-PGAAGS		153
QY	115	GAGAGSGAGAS	GAGSG	GGVGV	PGVGV	PGKGV	PGVG	-PGVG	-PGVG	-PGA		170
Db	154	GAGAGSGAGAS	GAGSG	GGVGV	PGVGV	PGKGV	PGVG	-PGVG	-PGVG	-PGA		213
QY	171	GAGSAGAGSGAGS	GAGSG	GGVGV	PGVGV	PGKGV	PGVG	-PGVG	-PGVG	-PGVG		227
Db	214	GAGSAGAGSGAGS	GAGSG	GGVGV	PGVGV	PGKGV	PGVG	-PGVG	-PGVG	-PGVG		273
QY	228	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVGV	PGKGV	PGVG	-PGVG	-PGVG			283
Db	274	VPGAAGSGAGAGS	GAGAGS	GGVGV	PGVGV	PGKGV	PGVG	-PGVG	-PGVG			333
QY	284	PGVG	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVGV	PGKGV	PGVG	-PGVG	-P		340
Db	334	PGVG	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVGV	PGKGV	PGVG	-PGVG	-P		393
QY	341	GGV	-PGVG	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVGV	PGKGV	PGVG	-PG		397
Db	394	GGV	-PGVG	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVGV	PGKGV	PGVG	-PG		453
QY	398	VG	-PGVG	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVGV	PGKGV	PGVG			454
Db	454	VG	-PGVG	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVGV	PGKGV	PGVG			513
QY	455	G	-PGVG	-PGVG	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVGV	PGKGV	PGVG		510
QY	514	GGV	-PGVG	-PGVG	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVGV	PGKGV	PGVG		573
Db	514	GGV	-PGVG	-PGVG	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVGV	PGKGV	PGVG		573
QY	511	VP	GGV	-PGVG	-PGVG	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVGV	PGVG		566
Db	574	VP	GGV	-PGVG	-PGVG	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVGV	PGVG		633
QY	567	PGKGV	PGVG	-PGVG	-PGVG	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVGV			622
Db	634	PGKGV	PGVG	-PGVG	-PGVG	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVGV			693
QY	623	GGV	PGKGV	PGVG	-PGVG	-PGVG	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVG		678
Db	694	GGV	PGKGV	PGVG	-PGVG	-PGVG	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVG		753
QY	679	GGV	PGKGV	PGVG	-PGVG	-PGVG	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVG		734
Db	754	GGV	PGKGV	PGVG	-PGVG	-PGVG	-PGAAGSGAGAGS	GAGAGS	GGVGV	PGVG		813
QY	735	GGV	PGKGV	PGVG	-PGVG	-PGVG	-PGAAGSGAGAGS					780
Db	814	GGV	PGKGV	PGVG	-PGVG	-PGVG	-PGAAGSGAGAGS					863

RESULT 2  
US-11-201-606-25  
Sequence 25, Application US/11201606  
Publication No. US20060029638A1  
GENERAL INFORMATION:  
APPLICANT: STEDRONSKY, Edwin R.  
APPLICANT: CARPELLO, Joseph  
TITLE OF INVENTION: Tissue Adhesive Using Synthetic  
TITLE OF INVENTION: CrossLinking  
NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEH, HOENACH, TEST, ALBRITTON & HERBERT  
STREET: Four Embarcadero Center, Suite 200  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/11/201,606  
FILING DATE: 2005-08-10  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/642,246  
FILING DATE: 1996-05-02  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: ROWLAND, Bertram I  
REGISTRATION NUMBER: 20015  
REFERENCE/DOCKET NUMBER: A61127-1/3IR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1002 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-11-201-606-25

Query Match	68.4%;	Score 2855;	DB 11;	Length 1002;
Best Local Similarity	65.6%;	Pred. No. 3.3e-189;		
Matches 668;	Conservative 4;	Mismatches 41;	Indels 306;	Gaps 57;

[illegible]

```

; INFORMATION FOR SEQ ID NO: 30:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 936 amino acids
;
; TYPE: amino acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
US-11-201-606-30

Query Match: 67.0%; Score 2795; DB 11; Length 936;
Best local similarity 73.9%; Pred. No. 4, 1e-185;
Matches 642; Conservative 4; Mismatches 93; Indels 130; Gaps 55;

```

[illegible]

RESULT 4  
US-11-201-606-34  
; Sequence 34, Application US/11201606  
; Publication No. US20060029638A1  
; GENERAL INFORMATION:  
; APPLICANT: STEDRONSKY, Edwin R.  
; APPLICANT: CAPPELLO, Joseph  
; TITLE OF INVENTION: Tissue Adhesive Using Synthetic  
; TITLE OF INVENTION: Crosslinking  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
; STREET: Four Embarcadero Center, Suite 200  
; City: San Francisco  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/11/201,606  
; FILING DATE: 2005-08-10  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/642,246  
; FILING DATE: 1996-05-02  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ROWLAND, Bertram I  
; REGISTRATION NUMBER: 20015  
; REFERENCE/DOCKET NUMBER: A61127-1/BIR  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-781-1989  
; TELEFAX: 415-398-3249  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 966 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-11-201-606-34

Query Match 65.7%; Score 2740; DB 11; Length 966;  
Best Local Similarity 72.5%; Pred. No. 2.5e-181;  
Matches 642; Conservative 4; Mismatches 104; Indels 136; Gaps 55;

QY 1 GAAGAGSAGAGSGVGP-----GTVGPGVPGKGVPGVPGV-PGVG-PGVG-P 48  
DB 94 GAAGAGSAGAGSGVGPFGFVRARRGVPGVPGVPGKGVPGVPGVPGVPGV 153  
QY 49 GAGAGSAGAGSGAGA-GSAGAGSGVGPVPGVPGKGVPGVPGVPGVPGV-PGV 106  
DB 154 GAGAGSAGAGSGVGPFGFVRARRGVPGVPGVPGKGVPGVPGVPGVPGV 210  
QY 107 G-PGAGAGSAGAGSGAGA-GSAGAGSGVGPVPGVPGKGVPGVPGVPGVPGV- 163  
DB 211 GVPGAGAGSAGAGSGVGPFGFVRARRGVPGVPGVPGKGVPGVPGVPGV- 267  
QY 164 PGVGP-PGAGAGSAGAGSGAGA-GSAGAGSGVGPVPGVPGVPGKGVPGVPGVPG 221  
DB 268 PGVGP-PGAGAGSAGAGSGVGPFGFVRARRGVPGVPGVPGKGVPGVPGVPG 324  
QY 222 VG-PGVG-PGAGAGSAGAGSGAGA-GSAGAGSGVGPVPGVPGKGVPGVPGVPGV 278  
DB 325 VGVGPVGP-PGAGAGSAGAGSGVGPFGFVRARRGVPGVPGVPGKGVPGVPGVPGV- 382  
QY 279 GPGVGP-PGVG-PGAGAGSAGAGSGAGA-GSAGAGSGVGPVPGVPGVPGKGVPGVPG 335

DB 383 -PGVGPVGPVPGAGAGSAGAGSGVGPFGFVRARRGVPGVPGVPGKGVPGV 441  
QY 336 PGVGPVGPV-PGVG-PGAGAGSAGAGSGAGA-GSAGAGSGVGPVPGVPGVPGKGV 392  
DB 442 V---PGVGPVGPVPGAGAGSAGAGSGVGPFGFVRARRGVPGVPGVPGKGV 498  
QY 393 GVPVGPVGPVGPV-PGVG-PGAGAGSAGAGSGAGA-GSAGAGSGVGPVPGVPGVPGK 449  
DB 499 GVGVP---PGVGPVGPVPGAGAGSAGAGSGVGPFGFVRARRGVPGVPGVPGVPGK 555  
QY 450 GVPVGPVGPVGPVGPV-PGVG-PGAGAGSAGAGSGAGA-GSAGAGSGVGPVPGVPGV 506  
DB 556 GVPVGPV---PGVGPVGPVPGAGAGSAGAGSGVGPFGFVRARRGVPGVPGVPGV 612  
QY 507 PGKGVPGVGPVGPVGPV-PGVG-PGAGAGSAGAGSGAGA-GSAGAGSGVGPVPGVPG 563  
DB 613 PGKGVPGVGPV---PGVGPVGPVPGAGAGSAGAGSGVGPFGFVRARRGVPGVPGV 669  
QY 564 VGVPGKVPVGPVGPVGPVGPV-PGVG-PGAGAGSAGAGSGAGA-GSAGAGSGVGPVPGV 620  
DB 670 VGVPGKVPVGPVGPV---PGVGPVGPVPGAGAGSAGAGSGVGPFGFVRARRGVPGV 726  
QY 621 VPGVGPVPGKVPVGPVGPVGPVGPV-PGVG-PGAGAGSAGAGS-----660  
DB 727 VPGVGPVPGKVPVGPVGPV---PGVGPVGPVPGAGAGSAGAGSGVGPFGFVRARRGV 783  
QY 661 -----GAGAGSAGAGSGVGP-----GVPV 682  
DB 784 GVGVPVGPVPGKVPVGPVGPVGPVGPVGPVPGVPGVPGAGAGSAGAGSGVGPFGFVRARRGV 843  
QY 683 GVGVPVGPVPGVGPVGPV-PGVG-PGVG-PGAGAGSAGAGSGAGA-GSAGAGSGVGP 737  
DB 844 GVGVPVGPVPGKVPVGPVGPVGPVGPVGPVPGVPGVPGAGAGSAGAGSGVGPFGFVRARRGV 903  
QY 738 GVGVPVGPVPGKVPVGPVGPV-PGVG-PGVG-PGVGPAGAGSAGAGS 780  
DB 904 GVGVPVGPVPGKVPVGPVGPVGPVGPVPGVPGVPGAGAGSAGAGSAGA 949

RESULT 5  
US-11-053-100-46  
; Sequence 46, Application US/11053100  
; Publication No. US2005025554A1  
; GENERAL INFORMATION:  
; APPLICANT: CHILKOTI, Ashutosh  
; TITLE OF INVENTION: FUSION PEPTIDES ISOLATABLE BY PHASE TRANSITION  
; FILE REFERENCE: 4176-101 CIP  
; CURRENT APPLICATION NUMBER: US/11/053,100  
; CURRENT FILING DATE: 2005-02-08  
; PRIOR APPLICATION NUMBER: US 09/812,382  
; PRIOR FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: US 60/190,659  
; PRIOR FILING DATE: 2000-03-20  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: Patentin version 3.3  
; SEQ ID NO 46  
; LENGTH: 1186  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1)-(1186)  
; OTHER INFORMATION: PET15b-SD3-ELP1-180-throm-Androgen receptor ligand binding domain  
US-11-053-100-46

Query Match 50.1%; Score 2091.5; DB 11; Length 1186;  
Best Local Similarity 59.2%; Pred. No. 8.4e-137;  
Matches 542; Conservative 2; Mismatches 223; Indels 149; Gaps 68;

QY 11 GSGVGPVGPVGPVGPVPGKGVPGVGPV-PGVG-PGVG-PGAGAGSAG-----GAG 59











```

; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)_(639)
; OTHER INFORMATION: pET32a-SD15-ELF4-120-BK-Insulin A peptide
US-11-053-100-25

Query Match      45.8%, Score 1910.5; DB 11, Length 639;
Best Local Similarity 59.7%; Pred. No. 1.3e-124;
Matches 491; Conservative 0; Mismatches 66; Indels 265; Gaps 67,

QY      11  GSGVGVPGVGVPGVGVPGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGSGAGAGSGAGAGS 66
DB      3   GPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGV----- 48
QY      67  GAGAGSGVGVPGVGVPGVGVPGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGSGAGAGSGA 122
DB      49  -----PGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGV----- 93
QY      123 GAGSGAGAGSGVGVPGVGVPGVGVPGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGSGAGA 178
DB      94  -----PGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGV----- 138
QY      179 GSGAGAGSGAGAGSGGVGVPGVGVPGVGVPGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGS 234
DB      139 -----PGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGV-- 183
QY      235 GAGAGSGAGAGSGAGAGSGGVGVPGVGVPGVGVPGKGVPGVG-PGVG-PGVG-PGVG-PGA 290
DB      184 -----PGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGV 226
QY      291 GAGSGAGAGSGAGAGSGAGAGSGGVGVPGVGVPGVGVPGKGVPGVG-PGVG-PGVG-PGVG 347
DB      227 GV-----PGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGV 267
QY      348 -PGAGAGSGAGAGSGAGAGSGAGAGSGGVGVPGVGVPGVGVPGKGVPGVG-PGVG-PGVG- 403
DB      268 VPGVGV-----PGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGV 308
QY      404 PGVG-PGAGAGSGAGAGSGAGAGSGAGAGSGGVGVPGVGVPGVGVPGKGVPGVG-PGVG-P 460
DB      309 PGVGVPGVGV-----PGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGV 349
QY      461 GVG-PGVG-PGAGAGSGAGAGSGAGAGSGAGAGSGGVGVPGVGVPGVGVPGKGVPGVG-PG 517
DB      350 GVGVPGVGVPGVGV-----PGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGV 390
QY      518 VG-PGVG-PGVG-PGAGAGSGAGAGSGAGAGSGAGAGSGGVGVPGVGVPGVGVPGKGVPGV 574
DB      391 VGVGVGVGVPGVGVGVGV-----PGVGVPGVGVPGVGVPGVGVPGVGVPGV 431
QY      575 G-PGVG-PGVG-PGVG-PGAGAGSGAGAGSGAGAGSGAGAGSGGVGVPGVGVPGVGVPGK 630
DB      432 GVPGVGVPGVGVPGVGVPGVGV-----PGVGVPGVGVPGVGVPGVGVPGV 472
QY      631 VPGVGV-PGVG-PGVG-PGVG-PGAGAGSGAGAGSGAGAGSGAGAGSGGVGVPGVGVPGVGV 686
DB      473 VPGVGVPGVGVPGVGVPGVGVPGVGV-----PGVGVPGVGVPGVGV 513
QY      687 PGKGVPGVGV-PGVG-PGVG-PGVG-PGAGAGSGAGAGSGAGAGSGAGAGSGGVGVPGVGV 742
DB      514 PGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGV-PGVGV-PGVGV-PGVGV-PGVGV 569
QY      743 GVGVPGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGSGAGAGS 780
DB      570 GVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGV 611

RESULT 13
US-11-053-100-29
; Sequence 29, Application US/11053100

```

Publication No. US20050255554A1  
GENERAL INFORMATION:  
APPLICANT: CHIKOTI, Ashutosh  
TITLE OF INVENTION: FUSION PEPTIDES ISOLATABLE BY PHASE TRANSITION  
FILE REFERENCE: 4176-101 CIP  
CURRENT APPLICATION NUMBER: US/11/053,100  
CURRENT FILING DATE: 2005-02-08  
PRIOR APPLICATION NUMBER: US 09/812,382  
PRIOR FILING DATE: 2001-03-20  
PRIOR APPLICATION NUMBER: US 60/190,659  
PRIOR FILING DATE: 2000-03-20  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 29  
LENGTH: 654  
TYPE: PRT  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: Synthetic Construct  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (1)-(654)  
OTHER INFORMATION: PET15b-ELP4-120-EK-T20 peptide  
US-11-053-100-29

Query Match 45.8%; Score 1910.5; DB 11; Length 654;  
Best Local Similarity 59.7%; Pred. No. 1,3e-124;  
Matches 491; Conservative 0; Mismatches 66; Indels 265; Gaps 67;

QY 11 GSGVGPVGVPVGVPVGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGSAGAGSAGAGS 66  
DB 3 GPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 48  
QY 67 GAGAGSVGPVGVPVGVPVGKGVPGVG-PGVG-PGVG-PGVG-PGVG-PGAGAGSAGAGS 122  
DB 49 -PGVGPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 93  
QY 123 GAGSAGAGSVGPVGVPVGVPVGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGSAGAG 178  
DB 94 -PGVGPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 138  
QY 179 GSGAGAGSAGAGSVGPVGVPVGVPVGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGS 234  
DB 139 -PGVGPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 183  
QY 235 GAGAGSAGAGSVGPVGVPVGVPVGKGVPGVG-PGVG-PGVG-PGVG-PGA 230  
DB 184 -PGVGPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 226  
QY 291 GAGSAGAGSAGAGSVGPVGVPVGVPVGKGVPGVG-PGVG-PGVG-PGVG-PGVG 347  
DB 227 GV-----PGVGPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 267  
QY 348 -PGAGAGSAGAGSAGAGSVGPVGVPVGVPVGKGVPGVG-PGVG-PGVG-PGVG- 403  
DB 268 VPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 308  
QY 404 PGVGPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 460  
DB 309 PGVGPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 349  
QY 461 GVG-PGVG-PGAGAGSAGAGSAGAGSVGPVGVPVGVPVGKGVPGVG-PG 517  
DB 350 GVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 390  
QY 518 VG-PGVG-PGVG-PGAGAGSAGAGSAGAGSVGPVGVPVGVPVGKGVPGVG 574  
DB 391 VGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 431  
QY 575 G-PGVG-PGVG-PGVG-PGAGAGSAGAGSAGAGSVGPVGVPVGVPVGKGV 630  
DB 432 GVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 472

QY 631 VPVG-PGVG-PGVG-PGAGAGSAGAGSAGAGSVGPVGVPVGVPVGVPVGVP 686  
DB 473 VPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 513  
QY 687 PGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGSAGAGSAGAGSVGPVGVPVGVP 742  
DB 514 PGVGPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 569  
QY 743 GVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 780  
DB 570 GVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 611

RESULT 14  
US-11-053-100-38  
Sequence 38, Application US/11053100  
Publication No. US20050255554A1  
GENERAL INFORMATION:  
APPLICANT: CHIKOTI, Ashutosh  
TITLE OF INVENTION: FUSION PEPTIDES ISOLATABLE BY PHASE TRANSITION  
FILE REFERENCE: 4176-101 CIP  
CURRENT APPLICATION NUMBER: US/11/053,100  
CURRENT FILING DATE: 2005-02-08  
PRIOR APPLICATION NUMBER: US 09/812,382  
PRIOR FILING DATE: 2001-03-20  
PRIOR APPLICATION NUMBER: US 60/190,659  
PRIOR FILING DATE: 2000-03-20  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 38  
LENGTH: 656  
TYPE: PRT  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: Synthetic Construct  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (1)-(656)  
OTHER INFORMATION: PET17b-ELP4-120-REV(Q/Y)-T20 peptide  
US-11-053-100-38

Query Match 45.7%; Score 1906.5; DB 11; Length 656;  
Best Local Similarity 59.8%; Pred. No. 2.5e-124;  
Matches 490; Conservative 0; Mismatches 65; Indels 265; Gaps 67;

QY 11 GSGVGPVGVPVGVPVGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGSAGAGSAGAGS 66  
DB 3 GPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 48  
QY 67 GAGAGSVGPVGVPVGVPVGKGVPGVG-PGVG-PGVG-PGVG-PGVG-PGAGAGSAGAGS 122  
DB 49 -PGVGPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 93  
QY 123 GAGSAGAGSVGPVGVPVGVPVGKGVPGVG-PGVG-PGVG-PGVG-PGAGAGSAGAG 178  
DB 94 -PGVGPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 138  
QY 179 GSGAGAGSAGAGSVGPVGVPVGVPVGKGVPGVG-PGVG-PGVG-PGVG-PGA 234  
DB 139 -PGVGPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 183  
QY 235 GAGAGSAGAGSAGAGSVGPVGVPVGVPVGKGVPGVG-PGVG-PGVG-PGVG-PGVG 290  
DB 184 -PGVGPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 226  
QY 291 GAGSAGAGSAGAGSAGAGSVGPVGVPVGVPVGKGVPGVG-PGVG-PGVG-PGVG-PGVG 347  
DB 227 GV-----PGVGPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 267  
QY 348 -PGAGAGSAGAGSAGAGSVGPVGVPVGVPVGKGVPGVG-PGVG-PGVG-PGVG- 403  
DB 268 VPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 308



**This Page Blank (uspto)**